

## ORIGINAL ARTICLE

# Dissociation of Neural Networks for Predisposition and for Training-Related Plasticity in Auditory-Motor Learning

Sibylle C. Herholz<sup>1,2,3,†</sup>, Emily B.J. Coffey<sup>1,2,†</sup>, Christo Pantev<sup>4</sup>, and Robert J. Zatorre<sup>1,2</sup>

<sup>1</sup>Montreal Neurological Institute, McGill University, 3801 rue University, Montreal, Quebec H2A 3B4, Canada,

<sup>2</sup>International Laboratory for Brain, Music and Sound Research (BRAMS), Montreal, Quebec H3C 3J7, Canada,

<sup>3</sup>German Center for Neurodegenerative Diseases (DZNE), Holbeinstr 13–15, Bonn 53175, Germany, and <sup>4</sup>Institute for Biomagnetism and Biosignalanalysis, University of Münster, Malmedyweg 15, Münster 48149, Germany

Address correspondence to Sibylle Herholz, German Center for Neurodegenerative Diseases (DZNE), Holbeinstr 13-15, 53175 Bonn, Germany.

Email: sibylle.herholz@dzne.de

<sup>†</sup>S.C.H. and E.B.J.C. contributed equally to this study.

## Abstract

Skill learning results in changes to brain function, but at the same time individuals strongly differ in their abilities to learn specific skills. Using a 6-week piano-training protocol and pre- and post-fMRI of melody perception and imagery in adults, we dissociate learning-related patterns of neural activity from pre-training activity that predicts learning rates. Fronto-parietal and cerebellar areas related to storage of newly learned auditory-motor associations increased their response following training; in contrast, pre-training activity in areas related to stimulus encoding and motor control, including right auditory cortex, hippocampus, and caudate nuclei, was predictive of subsequent learning rate. We discuss the implications of these results for models of perceptual and of motor learning. These findings highlight the importance of considering individual predisposition in plasticity research and applications.

**Key words:** auditory cognition, auditory imagery, fMRI, music, neuroplasticity

## Introduction

Practice of skills affects brain networks that are involved in both basic and higher-order task-related cognition (Jäncke 2009; May 2011; Zatorre, Fields et al. 2012). Musical training has emerged as a valuable framework to study the effects of learning complex tasks on the human brain (Münste et al. 2002; Zatorre et al. 2007; Herholz and Zatorre 2012) because it affects both auditory perception (Shahin et al. 2003; Bosnyak et al. 2004; Fujioka et al. 2004) and higher-order cognition such as auditory imagery (Herholz et al. 2008). But is it only practice that makes perfect? Individuals clearly differ in their ability to learn complex tasks,

and individual differences in brain function and structure, including in the auditory cortices, can predict learning rates in various tasks (Wong et al. 2008; Zatorre, Delhommeau, et al. 2012; Zatorre 2013). Neurophysiological measures of auditory perception and imagery show considerable individual variability (Schneider et al. 2002; Daselaar et al. 2010; Herholz et al. 2012), but individual predictors of complex multimodal learning are poorly understood. An important goal of our study was thus to identify any preexisting neural markers that predict learning outcome.

Multisensory integration during practice enhances training-related changes in sensory and association cortical areas during

auditory cognition (Lappe et al. 2008; Paraskevopoulos et al. 2012), and training results in stronger auditory-motor synchronization or co-activation during auditory and motor tasks (Bangert and Altenmüller 2003; D'Ausilio et al. 2006; Lahav et al. 2007). Musical experience also modulates performance and brain activity in auditory imagery tasks, as shown in cross-sectional studies comparing experts and novices (Aleman et al. 2000; Herholz et al. 2008). Auditory perception and imagery engage partly overlapping and partly distinct cortical networks (Herholz et al. 2012) and have been described as parallel processes, such that imagery represents the “offline” use of the top-down model of expected outcomes from perception and action (Grush 2004; Rauschecker and Scott 2009). Our second goal was therefore to test whether auditory-motor training would affect neural activity not only during perception but also during imagery, and to what extent changes for perceptual and more abstract cognitive tasks overlap.

Both exact encoding of incoming sensory information as well as predictions of upcoming sensory events and of outcomes of motor actions are important for learning. In the framework of predictive coding (Friston 2005), top-down internal models of rules and regularities generate expectations about sensory events and are modulated by error signals with the overall goal to minimize prediction error. Thus, bottom-up or forward information about sensory input is compared with predictions generated from higher-order levels of processing. Importantly, while forward input is relatively stable, learning occurs through the adaptation of the higher-order, backward, or top-down connections (Friston 2005). Based on these considerations, we expected that predisposition for learning versus experience-dependent plastic changes might dissociate at the level of brain networks.

Here, we investigated the effects of piano training using a longitudinal design that enabled us 1) to observe the causal influence of training on brain activity under naturalistic but controlled conditions and 2) to determine individual predictors of learning, within the same individuals. To test the effects of multisensory training on higher-order auditory cognition, we used a basic music perception task and a more demanding musical imagery task, which we administered while collecting BOLD fMRI data before and after 6 weeks of systematic piano training. We expected task-specific changes due to musical training in cortical auditory and motor areas, and comparable changes in the networks for auditory perception and imagery. Regarding individual predisposition, we hypothesized that pre-training levels of activity in the auditory-motor network would predict the subsequent performance during piano training. While both neuroplastic and predictive findings were expected in task-relevant brain areas, we aimed to dissociate specific components of these networks that either change through training or are predictive of learning.

## Methods

### Participants

Fifteen healthy, right-handed young adults (aged 20–34 years, average 25.6 years; 7 male) enrolled in the study. Fourteen participants completed the study. One person (female) dropped out before training due to reasons unrelated to the study. Participants were selected based on lack of musical background, as assessed by an online version of the Montreal Music History Questionnaire (Coffey et al. 2011), MRI compatibility, availability at the time of the study, and personal commitment. None of the participants had >2 years of formal musical training, none were currently engaging in active music making, and none had previously received any training on a keyboard instrument. All participants were

native English speakers or bilingual (English and French), had grown up in the United States of America or Canada and were familiar with the melodies used as stimulus material. Participants provided informed consent before enrolling. The ethics review board of the Montreal Neurological Institute, McGill University, approved the protocol.

### Materials

Twelve familiar melodies including Christmas carols, nursery rhymes, and pop songs were used as stimulus material. The melodies were piloted for familiarity on a different group of 10 Canadian and American participants that was comparable to the study group with respect to age and musical background. From those found to be universally familiar, we selected melodies that were relatively easy to play on the piano and that yielded above-chance performance in the imagery task (see below) in pilot participants. Stimuli were presented in piano timbre in the fMRI tasks (see below) and during training. Melodies were created as MIDI stimuli first (grand piano timbre, Anvil Studio, Willow Software) and then converted to wav-files (stereo, 44 100 Hz sampling rate, plug-in to Winamp). The headphones used in the fMRI experiment (S14, Sensimetrics Corporation) have a flat response in the range of 100 Hz to 8 kHz. Melodies were transposed as needed so that the training keyboard covered their tonal range and so that they were relatively easy to play. The tempo of the melodies varied, and for each melody, it was similar to popular recordings. The key and tempo of each melody was the same for the piano training and for the fMRI tasks. For a control condition in the fMRI experiment (see section “Musical cognition tasks” below), we prepared permuted tone sequences based on the familiar melodies that were physically comparable to the stimuli used in the imagery task but were unrecognizable and sounded unfamiliar.

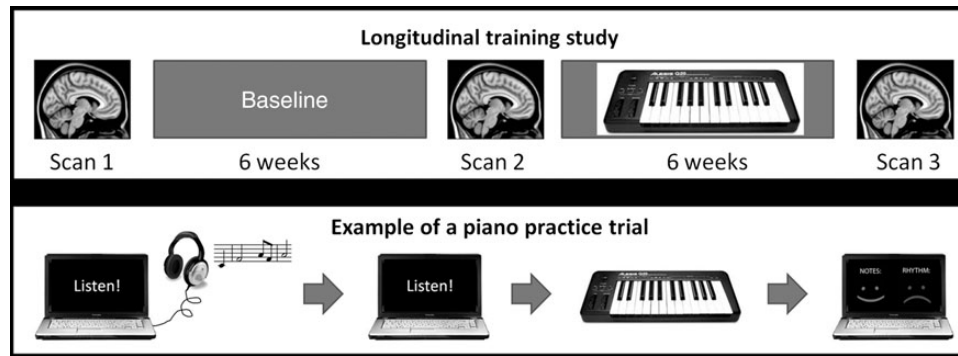
### Procedures

#### Longitudinal Study Design

We used a within-subject design with baseline and training periods, in which subjects were tested at 3 time points spaced 6 weeks apart (Fig. 1, top panel). Scans 1 and 2 took place before and after a baseline period of 6 weeks, during which no training occurred. Immediately after the second MRI session, participants began the 6-week piano training period, after which Scan 3 took place. This design allowed us to distinguish the effects of training from unrelated changes, as well as to determine individual predictors of learning (Thomas et al. 2009). For one subject, the baseline period had to be extended to 8 weeks due to unforeseen scheduling problems. In all other cases, scans were scheduled 6 weeks  $\pm$  2 days apart.

#### Piano Training Protocol

Participants took part in a 6-week piano training in home- and lab-based practice sessions. They took an electronic keyboard home that was connected to an online training system implemented with Presentation Software (Neurobehavioral Systems, Inc.). Participants learned to play simple tunes on the piano in practice sessions of 30-min' duration each, 5 days per week (30 sessions total), following a custom training curriculum. The piano-training protocol was created for the study and piloted for progression of difficulty on several additional subjects who were not included in the study. The training focused on the auditory and motor domains by having subjects listen to a short melody and repeat it on the piano on each trial. During the first



**Figure 1.** Illustration of the longitudinal within-subject repeated-measures design (upper panel) and illustration of an example trial of the piano training (lower panel). See text for details on the design and on the training program.

4 weeks (20 sessions), participants learned to play simple tone sequences that were composed for the study. Exercises were grouped in levels that focused on a specific skill to be taught, ranging from 3-tone isochronous sequences to be played by one hand only, to more complex rhythmic sequences that involved both hands. During Weeks 5 and 6 (10 sessions), participants learned to play 6 familiar melodies. These melodies were partitioned into smaller fragments that were successively taught and then put together to form the complete melodic phrase. Two randomly assigned subgroups learned to play different halves of the set of 12 melodies used in the functional paradigm (6 melodies per participant, counterbalanced across subjects with approximately equal difficulty and song genres in each subgroup). This allowed us to later test material-specific and generalized effects of training.

At the beginning of each level, participants saw a slide that explained what the goal of this level was, which hand they would be using, which keys on the piano they would need, and where they should place their hands. On each practice trial (Fig. 1, bottom panel), participants first listened to a tone sequence (template melody), cued by the word “listen” on the screen. Then, following the cue “practice,” a graphic appeared to indicate the piano key of the first tone. We included this starting reference because we observed that curriculum pilot subjects did not form sufficiently strong direct associations between the tones and key position during this short training period. However, they received no further visual or verbal information on what to play, and thus, training was mainly by ear. Participants repeated the melody to their best ability. Once participants had struck the same number of keys as in the template melody, they received visual feedback about the correctness of the keys and the rhythmic timing. Positive and negative feedback was given in form of a smiling or sad face for the tones and a smiling, neutral, or sad face for the rhythm. Participants had to repeat the exercise if they received negative feedback regarding tones, rhythm or both. In the first part of the training, each exercise could only be repeated up to 3 times. After the third failure, the order of exercises within the corresponding level was shuffled and a different melody was presented on the next trial. This prevented participants from getting stuck with one particular melody, and limited frustration. Participants passed a level if they correctly played all melodies (within 3 attempts per melody) in one run.

In the last 2 weeks of training (10 sessions), participants practiced familiar melodies, which were also used in the fMRI experiment. The rules regarding feedback were the same as in the first 4 weeks, but we omitted the shuffling of exercises since they were meant to successively make up the melody and scrambling

would have confused participants. Here, exercises had to be completed in order. At the end of each level, the complete melody had to be performed correctly 3 times before passing on to the next level (i.e., the next melody). If participants completed all melodies before the end of the 10 sessions, they went through the melodies a second time. After this, they practiced filler melodies that were from similar genres and of similar difficulty, but that were not used in the fMRI protocols. In the last training session that took place immediately before the last MRI session (see below), participants reviewed all 6 familiar melodies. Each melody was practiced for 5 min, to ensure that all participants had the same exposure to all trained melodies immediately before the scanning.

Participants practiced on 25-key midi keyboards (Q25, Alesis) interfacing with a custom training program created with Presentation (Neurobehavioral Systems, Inc.). Participants came to the lab once per week for a supervised training session, so that we could monitor the training progress personally and to answer questions. Participants practiced at home without personal supervision for the remaining 4 sessions per week. Training sessions were limited to 30 min by the software. Output files containing detailed information about the performance in the session were automatically uploaded to a lab-based ftp-server at the end of each session, so we were able to remotely monitor progress and compliance with the program and to remind participants about their schedule if necessary. Participants were instructed to do only one 30-min session per day and generally complied with this rule, but in rare cases, we allowed participants to do 2 sessions on a single day in order to accommodate their individual schedules for a particular week.

### MRI Scanning Session

Functional and structural MRI scans were collected at each of the 3 measurement time points in the study (Scans 1 to 3). The duration of each session was ~2 h. Scans took place at the McConnell Brain Imaging Center at the Montreal Neurological Institute on a 3 Tesla MR scanner with a 32-channel head coil (Siemens Trio, Erlangen, Germany). Auditory stimuli were delivered via MRI-compatible headphones (S14, Sensimetrics Corp.) with foam inserts placed inside the ear canal. Stimuli were delivered at a volume deemed comfortable by each participant. Visual stimuli were presented via back projection on a screen placed at the end of the MRI bore. Responses were recorded via an MRI-compatible button box.

### Musical Cognition Tasks

Participants performed musical cognition tasks during 2 runs, with 4 conditions that involved judging the correctness of the

last tone of a familiar melody (Listen), imagining part of the melody and judging if a final tone correctly completed the imagined tune (Imagine), listening to the random tone sequences and pressing a response key but without an auditory cognition task (Random), or resting in silence (Baseline). All conditions are illustrated in Figure 2. Briefly, on each trial, first a visual cue for the condition and the title of the melody (in Listen and Imagine conditions) were presented for 1 s, followed by the auditory stimulus or silence, immediately followed by the scan acquisition. The duration of auditory stimuli was on average 10.1 s. Participants were instructed to respond (Imagine, Listen, and Random conditions) during the scanner noise following the auditory stimulus or silence. Conditions were presented in blocks of 12 trials each for listen, imagine, and random conditions interspersed by blocks of 4 trials of the baseline condition. The order of the stimuli within the blocks was pseudo-randomized for each block. The order of the blocks was counterbalanced across subjects and different for the 2 runs for each subject. In total, 48 trials of each condition were presented.

**Listen condition.** On each listen trial, a melodic excerpt as described under Materials was presented. The last tone of the excerpt was incorrect in half of the trials. Participants had to indicate their judgment of correctness via button press. Incorrect tones were always in key, and excerpts ended before the end of the melodic phrase. Thus, participants could not use harmonic cues for their decision.

**Imagine condition.** The same familiar melodies as in the Listen condition were presented. However, instead of the full melody, only an initial segment of melody was presented, followed by a silent gap of at least 6 s. The last tone of the excerpt (as in Listen condition) was then presented again at the same time point when it would occur in the original melody. Participants' task was to imagine the continuation of the melody during the silent gap and judge the correctness of the last tone. This task was adapted for fMRI from a previous MEG study on mental imagery of music (Herholz et al. 2008). Since participants can only judge the last tone if they have correctly imagined the preceding part of the melody, it provides an objective measure of musical imagery.

**Random condition.** We included the random tone condition to control for the acoustic input that occurred in the imagery condition. These stimuli had the same physical parameters and silent gaps as in the Imagine condition, but the melody had been replaced by randomly scrambled versions containing the same pitch and duration of tones. Two scrambled versions were used for each melody in order to minimize an increase in familiarity

over the scanning session. Furthermore, 2 independent raters screened the scrambled melodies to exclude the possibility that any evoked a different familiar melody. This control was intended to reduce the possibility of spontaneous imagery of familiar songs. The same last tones as in the imagery and listen condition were presented, but since the initial tones were scrambled, they were not meaningful. No judgments regarding the tone were required from the participants in this condition, but in order to also control for the motor output of imagery and listen conditions, participants were instructed to press a button after the presentation of the last tone in this condition too.

**Baseline condition.** In the baseline control condition, no auditory input was presented. Participants were instructed to rest with eyes open during these trials, and not to perform any button presses.

### Scanning Parameters

We used a sparse sampling paradigm for functional scanning (Belin et al. 1999; Hall et al. 1999), that is, the volume acquisition took place after the presentation of the tones, and the stimuli were presented during the silent periods in between acquisitions. The timing of the volume acquisition was optimized to pick up listen- and imagery-related activity in these conditions and to avoid picking up activity related to processing of the initial tones in the imagery condition. We included the random tone condition to enable removal of any residual influences of the auditory stimulation in the analysis. We recorded EPI images covering the whole head (voxel size 3.4 mm<sup>3</sup>, 42 slices, TE 30 ms, TR 15 000 ms) immediately after the last tone was presented (Listen, Imagery, and Random conditions) or after an equivalent lapse of time (Baseline condition) (See Fig. 2). Between the first and second functional imaging run, we recorded anatomical T1-weighted images (MPRAGE, voxel size 1 mm<sup>3</sup>).

### Data Analysis

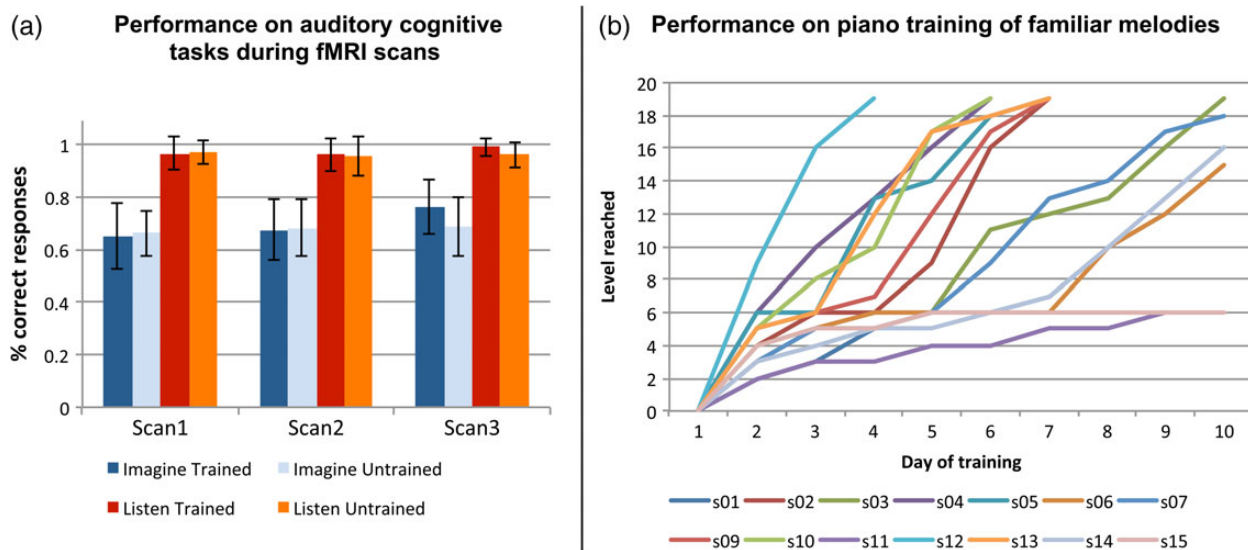
#### Behavioral Data Analysis

During training, participants progressed to a new curriculum level after successfully completing each set of exercises. We recorded the level reached at the end of each session and calculated a linear trend line for each subject. The slope of these were taken as a measure of the relative speed with which subjects progressed from the first to the nineteenth levels that made up the second part of the piano training (during which participants learned to play familiar melodies; see Fig. 3b).



**Figure 2.** Illustration of example trials of all conditions in the functional MRI paradigm. In these examples, the last tone for the Listen condition is correct, and incorrect for the Imagine condition. In the Random condition, no judgment of the last tone had to be given. Still, tones were the same 2 possible tones as in the respective Imagine condition. In the baseline condition, no acoustic stimuli were presented. The EPI BOLD sequence was run immediately after the last tone or after lapse of time (baseline condition), in order to optimally pick up activity related to the task. Participants gave their responses during the EPI scanning sequence.





**Figure 3.** (a) Group average performance on the imagery and listen tasks in the fMRI paradigm, for trained and untrained melodies and on each scan. Error bars represent standard deviations. (b) Individual performance curves during the melody part of the piano training. Numbers on the y-axis represent the number of melodies that were successfully learned.

Although power law fits are frequently used to model learning data, this approach can often be problematic (Clauzet et al. 2009). Instead, we used linear modeling because it requires few assumptions and provided excellent fits (see below). Furthermore, our aim was to capture relative differences between individuals as input for other analyses, rather than to describe the shape of each participant's learning curve.

For the auditory cognitive tasks in the fMRI session, mean correct responses were analyzed in repeated-measures analyses of variance. Alpha level was 0.05 for all analyses.

#### fMRI Data Analysis

Analyses were performed using FSL software (fMRIB, Oxford, UK; RRID:birnlex\_2067 and RRID:nif-0000-00305) (Smith et al. 2004; Jenkinson et al. 2012). For preprocessing, images were motion-corrected and spatially smoothed (5 mm FWHM). Individual fMRI data were registered to the individual's T1-weighted anatomical images (3-parameter linear transformation) and registered to MNI standard space for third-level analyses (12-parameter linear transformation). Task-related BOLD responses of each run were analyzed within the GLM (FEAT; Beckmann et al. 2003; Woolrich et al. 2004; Woolrich 2008), including all 4 conditions in the model (Listen, Imagine, Random, Baseline). For each individual scan, contrast images were computed for Listen vs. Random and Imagine vs. Random to assess basic task-related activity and changes of activity due to training. To also assess stimulus-specific training effects, we computed the following contrasts between each individual's trained and untrained melodies in a separate analysis: Listen trained vs. Listen untrained and Imagine trained vs. Imagine untrained. In this analysis, the Random and Baseline conditions were also included in the model but were not used in contrasts. For both analyses, the second analysis step was to combine contrast images of each run within one scan for each individual in a fixed-effects model in subject space. On the third level, comparisons across runs and between subjects were made in a random effects model in MNI space (FLAME1 in FSL). For correction of multiple comparisons, we applied cluster-corrected thresholds ( $z > 2.3$ ,  $P < 0.05$  cluster-corrected) as implemented in FSL, using a Z statistic

threshold to define contiguous clusters, followed by estimation of significance level of each cluster based on the cluster probability threshold (Worsley 2001). Tables show significant local peaks for each cluster that were located in gray matter according to the Harvard-Oxford cortical and subcortical atlases and cerebellar atlas (Diedrichsen et al. 2009) implemented in FSL. Training effects were assessed in the comparison of Scans 2 and 3. We also compared activations between Scans 1 and 2 as a baseline during which no changes in task-related activity were expected. To identify predictors of subsequent training success, we performed regression analyses of task-related activity (Imagine vs. Random, Listen vs. Random) during Scan 2 using the melody-learning rate measure as a regressor. This analysis would therefore reveal if activity patterns prior to learning were correlated with individual rates of subsequent learning.

## Results

### Behavioral Data

The individual learning trajectories of the 14 participants in the melody training phase are given in Figure 2b. Although the variability of learning rates was considerable, by the end of the training phase, all subjects had successfully learned to play at least the 6 melodies assigned to them for training. The choice of linear slopes as a measure of relative training rate appears justified given the excellent fit in the vast majority of subjects (mean  $r^2 = 0.92$ ,  $SD = 0.11$ ; 13 of 14  $r^2 > 0.88$ ).

Average performance during scanning in the Imagery and Listen conditions, split for trained and untrained melodies (balanced across subjects), is shown in Figure 2a. Participants performed above chance in both the Listen and Imagery conditions at all time points (all  $P < 0.05$ , one-sample t-tests, Bonferroni corrected for multiple comparisons). Thus, participants were able to accurately imagine the songs as evidenced by their above-chance performance in judging correct or incorrect continuations of the melodies following the imagery interval. Performance in the Listen condition was better than the Imagery condition in all subjects, and in the Listen condition, we observed

a clear ceiling effect, as expected since this is a comparatively simple task even for nonmusicians. Nonparametric tests comparing the 2 conditions showed significant differences at each time point (Wilcoxon tests, all  $P < 0.001$ ). Due to violation of the normality assumption for the Listen condition, we computed a repeated-measures ANOVA only for the Imagery condition, with factors time point (Scans 1, 2, and 3) and training (trained vs. untrained melodies), and found a main effect of time point ( $F_{2,24} = 4.58$ ,  $P = 0.021$ ) but no other significant effects. A one-way ANOVA on the factor time point with planned comparisons revealed no significant change during baseline (Scan 1 vs. Scan2), and a significant increase after compared with before training (Scan 2 < Scan 3,  $t_{12} = 2.02$ ,  $P = 0.027$ , one-tailed). Thus, training resulted in an improvement of task performance on the imagery task.

### Functional Imaging Data

To establish a basis for interpretation, we first identified basic networks of melody imagery and perception. At baseline (Scan 1, before any training took place), listening to familiar melodies (Listen > Random) strongly activated bilateral primary and secondary auditory cortices, bilateral thalamus, and parts of the motor network (caudate, cerebellum lobule VI and crus I, left precentral gyrus) as expected. Imagery of familiar melodies, controlling for auditory input (Imagery > Random), resulted in activity in secondary auditory cortices, superior parietal and inferior frontal cortices, as well as the motor network comprising left precentral gyrus, SMA, putamen, and cerebellum, lobule VI (see [Supplementary Table 1](#)). Building on this basic analysis of the imagery and music listening networks, we then focused on our 2 main questions: the effects of 6 weeks of piano training on auditory perception and imagery, and preexisting individual predictors of subsequent learning.

#### Training Effects

We analyzed changes in task-related activation across the baseline period (Scan 2 vs. Scan 1) and across the training period (Scan 3 vs. Scan 2). Across the baseline period, we found no significant changes other than a decrease of activity in medial frontal and paracingulate cortex during imagery ([Supplementary Table 2](#)). In contrast, for training-related effects (Scan 3 > Scan 2), we found similar increases over time for both the imagery and for the perception tasks. For all melodies, including the ones that were not trained, left dorsal premotor cortex activity increased after training for both task contrasts (Imagery > Random, Listen > Random), with a partial overlap of the significant clusters as confirmed in a conjunction analysis (intersection of significant clusters). For the contrast Imagery > Random, we also found a training-related increase in a cluster comprising supramarginal and postcentral gyri (Fig. 4a and [Supplementary Table 2](#)). In order to assess specific training effects, we analyzed the changes over time for trained compared with untrained melodies for each task (Imagine trained > Imagine untrained, Listen trained > Listen untrained). The pattern of changes was similar for both tasks: for trained melodies compared with untrained melodies, we found training-related increases of activity in premotor and dorsolateral prefrontal cortex, and in bilateral posterior parietal cortex, including intraparietal sulcus. Again, the significant clusters in both task conditions overlapped. Additionally, we observed training-related increases of activity in bilateral cerebellum (lobule VI) and training-related decreases in lateral occipital cortex for the contrast Imagery trained > Imagery untrained ([Supplementary](#)

[Table 2](#)). No changes were observed in subcortical regions or in auditory regions.

In summary, we observed both general and material-specific training effects caused by the 6-week piano training. Changes that occurred for all melodies were limited to left motor/premotor cortex close to the representation of the right hand, whereas more extensive changes were found for trained melodies compared with untrained melodies, encompassing left dorsal premotor and prefrontal cortex and bilateral posterior parietal cortex, with more extensive activations on the left. As predicted, training resulted in improvement of task performance on the imagery task. The functional imaging data showed very few changes during the baseline period, as expected, whereas after the training, there were many relevant changes in activity. Both findings point to the success and specificity of the training protocol and validate our tasks.

#### Predictors of Subsequent Learning Rates

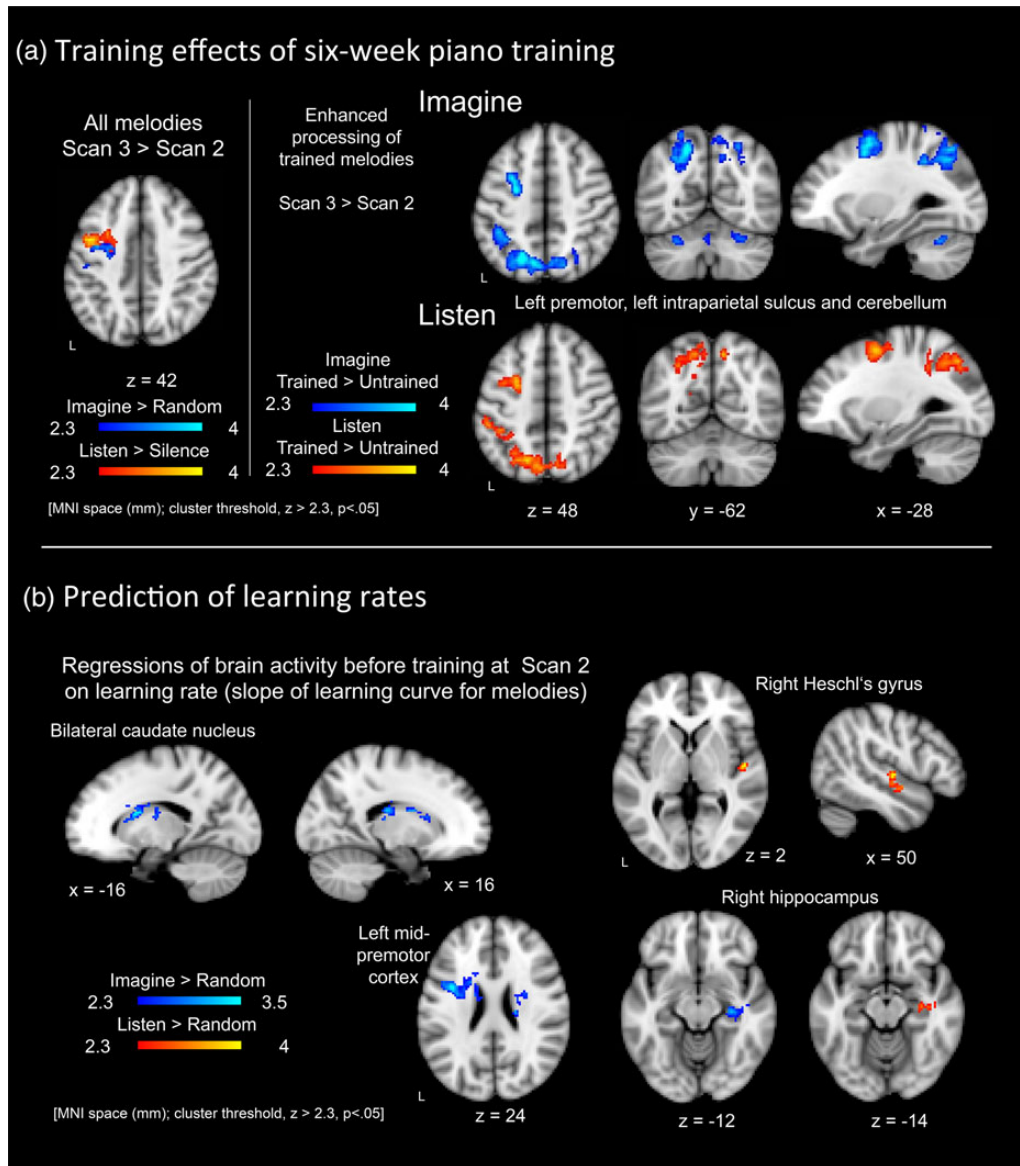
Individual learning rates during the melody training (last 2 weeks) were predicted by stronger activity in multiple areas pre-training, during both listening and imagining (Fig. 4b and [Supplementary Table 3](#)). We used the learning rates associated with the 2-week period of melody training, since this part of the training was most related to the melody perception and imagery tasks used in fMRI. For the Listen condition (Listen > Random), more activity in right auditory cortex (lateral portion of Heschl's gyrus [HG]) and right hippocampus predicted higher subsequent learning rates. For the Imagine condition (Imagine > Random), more activity in bilateral caudate (extending into thalamus), left mid-premotor cortex, and right hippocampus predicted higher learning rates. Less activity in several brain regions also predicted higher learning rates: less activity in medial frontal areas and frontal pole for both Listen and Imagine conditions, and less activity in occipital and precuneus cortex during Imagine ([Supplementary Table 3](#)). These analyses were designed to capture the extent to which variance in behavior measured at a later time point could be explained by activity patterns acquired at an earlier time point; the use of the term prediction is made in this context, rather than in the context of "out-of-sample" prediction. It remains to be seen how well activity measured in regions identified in the present study would transfer to a separate population; but logically, it is necessary to demonstrate prediction in the current context before being able to address this question in future research.

## Discussion

The key findings of this study are the clear evidence for both training-related neuroplasticity and neural predisposition for auditory-motor learning, with a dissociation of their respective brain substrates.

#### Training-Related Plasticity

We showed that auditory-motor training enhances activation of areas involved in motor preparation and sensorimotor integration during perception and imagery of familiar melodies, including premotor and posterior parietal areas in both conditions, and cerebellar hemispheres during imagery. Our findings extend previous findings of training-related plasticity and auditory-motor coactivation ([D'Ausilio et al. 2006](#); [Lahav et al. 2007](#)) to covert mental tasks. Going beyond previous cross-sectional studies on expertise effects in mental imagery ([Aleman et al. 2000](#); [Herholz et al. 2008](#)), we demonstrate causal effects of active auditory-motor training on neural correlates of imagery. Put more generally,



**Figure 4.** (a) Functional activation changes due to the 6-week piano training for imagery and listen conditions, and for all melodies (left panel) and specifically for trained melodies (right panel, assignment counterbalanced across subjects). Specific training-related changes are seen in left premotor and bilateral intraparietal sulcus for both conditions, and in cerebellar lobule IV for imagery. (b) Functional activity during imagery and listen conditions at Scan 2 that predicts subsequent training rates (slopes of learning curves) as estimated by linear regression. Individual differences in right hippocampus and right Heschl's gyrus were predictive during the listen condition, whereas clusters in right hippocampus, bilateral caudate and left mid-prefrontal cortex were predictive during the imagery condition.

the results show that training-related modulation of sensory-motor networks is also relevant for tasks that are more abstract in nature (i.e., without an overt sensory or motor component).

Premotor cortex can be subdivided into ventral and dorsal subregions that have been characterized as supporting direct versus abstract mapping of stimulus–action relationships (Petrides 1985; Fogassi et al. 2001; Kohler et al. 2002; Pizzamiglio et al. 2005; Hoshi and Tanji 2006; Zatorre et al. 2007). The complexity and large tonal range of the musical material in our study required computations of complex auditory-motor response selections, and the changes in dorsal premotor cortex are consistent with its role for abstract auditory-motor mapping (Zatorre et al. 2007; Chen et al. 2008a; 2008b; Chen et al. 2012). Posterior parietal regions support mental transformations of acoustic or visual information into motor representations (Stewart et al. 2003; Warren

et al. 2005; Zatorre et al. 2007; Brown et al. 2013) and mental musical transformations (Foster and Zatorre 2010; Foster et al. 2013), consistent with our task demands. Another parietal subregion, left supramarginal gyrus, that showed training-related increases for the imagery task has also previously been implicated in motor imagery (Hanakawa et al. 2008). Finally, the modulation of cerebellum activity after piano training is likely related to its implication in sensorimotor integration and formation of internal models of movements for motor sequence learning (Penhune and Steele 2012).

Changes in activity in the fronto-parietal network might be explained by more focused attention to the stimuli after training. If the effect of training had resulted in some generalized attentional enhancement then one would expect it to apply to all melodies, not only to the specific melodies trained. However, the

principal training-related effects are obtained in the contrast of trained versus untrained melodies. One possible explanation is that training altered the top-down control of attention specifically or to a larger extent for the trained melodies compared with the untrained melodies.

The pattern of parallel changes in premotor and parietal regions for both perception and imagery tasks is also consistent with their role in representing complex internal models, predictions, and transformations in auditory-motor learning. During explicit, initial stages of motor learning, prefrontal and parietal cortices and lateral cerebellar association areas are thought to store representations of learned sequences (Doyon and Benali 2005; Penhune and Steele 2012), and these areas would thus be expected to change their activity due to training. In perceptual learning, the predictive coding model suggests that top-down representations about rules and regularities are adapted through learning (Friston 2005). Consistently, structures important for adaptation of top-down processing, possibly including top-down control of attention, including premotor and parietal cortices and cerebellar hemispheres, showed training-related changes of activity in our study. Across both the perceptual and motor-learning models, the top-down predictions and internal models that are assumed to change through learning are also the most important conceptual parallel of imagery and perception/action in the online-offline model of forward models and efferent copies (Grush 2004; Rauschecker and Scott 2009). The regions of change and the parallel findings of training-induced changes across both tasks are in agreement with both theories.

### Predisposition for Learning

We identified several cortical and subcortical regions whose activity pre-training predicted subsequent learning rates on the piano training, and which were distinct from the areas that showed training-related changes. These included right HG, left mid-premotor cortex, bilateral caudate nucleus, and right hippocampus. The central roles of these regions in the context of auditory-motor training and auditory cognition can be attributed to encoding of stimuli (HG and hippocampus) and aspects of motor control (premotor cortex and caudate nuclei).

Lateral HG and adjacent regions, particularly of the right hemisphere, are known to be involved in fine-grained spectral processing and pitch discrimination from a large number of prior studies using both lesion and functional imaging approaches (Zatorre 1988; Johnsrude et al. 2000; Zatorre and Belin 2001; Patterson et al. 2002; Fujioka et al. 2003; Krumbholz et al. 2003; Penagos et al. 2004). Our result extends previous findings of HG function and structure as predictors of purely auditory learning (Wong et al. 2008; Zatorre, Delhommeau, et al. 2012) to more complex auditory-motor learning. Here, the predictive role of right auditory cortex most likely reflects enhanced encoding of pitch relationships, such that subsequent mapping of pitch to motor sequences is facilitated. While the perception and imagine tasks did not require episodic memory encoding, both tasks required retrieval of familiar melodies from long-term memory and maintenance in working memory. Right hippocampus activity has been related to successful melodic memory retrieval (Watanabe et al. 2008), and joint prefrontal and hippocampal activity are associated with high-load working memory tasks (Finn et al. 2010). Also, our piano training required participants to store a melody on each trial for immediate reproduction, and to compare template to outcome in order to correct errors regarding pitches and rhythm. It is notable that the predictive relationship with the hippocampus is found in 2 independent task conditions.

Enhanced processing of melodies in hippocampus during pre-learning perception and imagery tasks might therefore be reflected in enhanced performance during training.

In contrast to the training-related changes in dorsal premotor cortex discussed earlier, activity in more ventral premotor areas was predictive of subsequent learning rates. Left mid-premotor/inferior frontal cortex is involved in both mental imagery (Herholz et al. 2012) and in auditory-motor mapping, even when no clear sound-action association is established yet (Chen et al. 2012), indicating the beneficial effect of preexisting basic abilities to directly map sounds to actions. Ventral premotor cortex may support the initial basic ability to map actions to sounds, whereas more training-specific, complex mapping (dorsal premotor) needs to be established through training.

The caudate nucleus is activated in motor imagery (Nedelko et al. 2012), supports encoding of motor associations, and chunks during motor learning (Penhune and Steele 2012) and is active during the associative phase of cognitive procedural learning (Hubert et al. 2007). It is also active in musicians during tonal working memory (Schulze et al. 2011). This role of the caudate in challenging motor and auditory cognitive tasks is consistent with our finding that the level of caudate recruitment during auditory cognition predicted learning rates in auditory-motor training. Furthermore, in models of motor learning, the dorsolateral striatum, including caudate, supports chunking and fine-tuning of movements throughout and is thus crucial for learning but not expected to change through training—quite the opposite: Fine-tuning requires continuous practice, even in experts such as musicians and athletes (Penhune and Steele 2012).

### Conclusion

Learning can be conceived as interplay of bottom-up and top-down processing. According to the predictive coding theory (Friston 2005), top-down predictions based on bottom-up sensory input are continuously refined through evaluation of the mismatch of prediction and experience. Our results show that neural substrates that might be considered most relevant for bottom-up or forward input, encoding and chunking of auditory and sensorimotor information, including auditory cortex, hippocampus, and caudate, are predictive of subsequent learning success, in line with the idea that the quality of the forward signals is crucial for learning but does not necessarily change through training, at least not in the short term. In turn, training modulated activity in regions that integrate sensory information (parietal regions) and that select the appropriate motor programs during training (dorsal premotor cortex) and that are thus responsible for adapting responses to achieve correct auditory-motor performance and minimize errors. Findings from visuo-motor learning (den Ouden et al. 2010) and models of speech learning (Hickok and Poeppel 2007; Rauschecker and Scott 2009) also propose an interplay of prediction and feedback signals between auditory, motor, and association areas. This interpretation also fits with the concept of an efference copy of expected perceptual or action outcomes as an “online” top-down model, with parallel “offline” use of this model in mental imagery (Grush 2004; Rauschecker and Scott 2009). Our data thus match these combined concepts both regarding regions of change and parallelism of changes across Imagery and Listen tasks.

Cross-sectional studies in domains of expertise, including musicianship, cannot rule out the influence of predisposition, but this problem is often phrased as if it were merely a relatively minor drawback that would be overcome with corresponding findings of neuroplasticity in longitudinal training studies. Here, we show that predisposition plays an important role for



auditory-motor learning that can be clearly distinguished from training-induced plasticity. This dissociation contributes to our understanding of how the initial state of the nervous system can influence both the behavioral outcome of learning and its associated neural features (Zatorre 2013). Our findings pertain to the debate about the relative influence of “nature or nurture,” but also have potential practical relevance for individualized medicine and education, where they could help create customized interventions. Different interventions might be selected for individuals based on their predisposition and needs. The extent to which individual differences in predisposition are themselves the outcome of plasticity due to previous experiences in other domains and/or to (epi)genetic variability remains an important topic for cognitive neuroscience in the future.

## Supplementary Material

Supplementary material can be found at: <http://www.cercor.oxfordjournals.org/>.

## Funding

This study was funded by operating grants to R.J.Z. from the Canadian Institutes of Health Research, and by infrastructure funding from the Canada Fund for Innovation. S.C.H. was supported by Deutsche Forschungsgemeinschaft (He6067/1-1 and 3-1). E.B.J.C. is supported by a Vanier Canada Graduate Scholarship.

## Notes

We thank Stefanie Scala for assistance with stimulus preparation and piloting, Virginia Penhune and Chris Steele for helpful comments on the study design and analysis, and Virginia Penhune for helpful comments on an earlier version of the manuscript. *Conflict of interest statement:* None declared.

## References

- Aleman A, Nieuwenstein MR, Bocker KB, de Haan EH. 2000. Music training and mental imagery ability. *Neuropsychologia*. 38:1664–1668.
- Bangert M, Altenmüller EO. 2003. Mapping perception to action in piano practice: a longitudinal DC-EEG study. *BMC Neurosci*. 4:26.
- Beckmann CF, Jenkinson M, Smith SM. 2003. General multilevel linear modeling for group analysis in fMRI. *Neuroimage*. 20:1052–1063.
- Belin P, Zatorre RJ, Hoge R, Evans AC, Pike B. 1999. Event-related fMRI of the auditory cortex. *Neuroimage*. 10:417–429.
- Bosnyak DJ, Eaton RA, Roberts LE. 2004. Distributed auditory cortical representations are modified when non-musicians are trained at pitch discrimination with 40Hz amplitude modulated tones. *Cereb Cortex*. 14:1088–1099.
- Brown RM, Chen JL, Hollinger A, Penhune VB, Palmer C, Zatorre RJ. 2013. Repetition suppression in auditory-motor regions to pitch and temporal structure in music. *J Cogn Neurosci*. 25:313–328.
- Chen JL, Penhune VB, Zatorre RJ. 2008a. Listening to musical rhythms recruits motor regions of the brain. *Cereb Cortex*. 18:2844–2854.
- Chen JL, Penhune VB, Zatorre RJ. 2008b. Moving on time: brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training. *J Cogn Neurosci*. 20:226–239.
- Chen JL, Rae C, Watkins KE. 2012. Learning to play a melody: an fMRI study examining the formation of auditory-motor associations. *Neuroimage*. 59:1200–1208.
- Clauset A, Shalizi CR, Newman ME. 2009. Power-law distributions in empirical data. *SIAM Rev*. 51:661–703.
- Coffey EB, Herholz SC, Scala S, Zatorre RJ. 2011. Montreal Music History Questionnaire: a tool for the assessment of music-related experience in music cognition research. *The Neurosciences and Music IV: Learning and Memory, Conference, Edinburgh, UK, June 9–12 2011*.
- Daselaar SM, Porat Y, Huijbers W, Pennartz CM. 2010. Modality-specific and modality-independent components of the human imagery system. *Neuroimage*. 52:677–685.
- D’Ausilio A, Altenmüller E, Olivetti Belardinelli M, Lotze M. 2006. Cross-modal plasticity of the motor cortex while listening to a rehearsed musical piece. *Eur J Neurosci*. 24:955–958.
- den Ouden HE, Daunizeau J, Roiser J, Friston KJ, Stephan KE. 2010. Striatal prediction error modulates cortical coupling. *J Neurosci*. 30:3210–3219.
- Diedrichsen Jr, Balsters JH, Flavell J, Cussans E, Ramnani N. 2009. A probabilistic MR Atlas of the human cerebellum. *Neuroimage*. 46:39–46.
- Doyon J, Benali H. 2005. Reorganization and plasticity in the adult brain during learning of motor skills. *Curr Opin Neurobiol*. 15:161–167.
- Finn AS, Sheridan MA, Kam CL, Hinshaw S, D’Esposito M. 2010. Longitudinal evidence for functional specialization of the neural circuit supporting working memory in the human brain. *J Neurosci*. 30:11062–11067.
- Fogassi L, Gallese V, Buccino G, Craighero L, Fadiga L, Rizzolatti G. 2001. Cortical mechanism for the visual guidance of hand grasping movements in the monkey: A reversible inactivation study. *Brain*. 124:571–586.
- Foster NE, Halpern AR, Zatorre RJ. 2013. Common parietal activation in musical mental transformations across pitch and time. *Neuroimage*. 75:27–35.
- Foster NE, Zatorre RJ. 2010. A role for the intraparietal sulcus in transforming musical pitch information. *Cereb Cortex*. 20:1350–1359.
- Friston K. 2005. A theory of cortical responses. *Philos Trans R Soc Lond B Biol Sci*. 360:815–836.
- Fujioka T, Ross B, Okamoto H, Takeshima Y, Kakigi R, Pantev C. 2003. Tonotopic representation of missing fundamental complex sounds in the human auditory cortex. *Eur J Neurosci*. 18:432–440.
- Fujioka T, Trainor LJ, Ross B, Kakigi R, Pantev C. 2004. Musical training enhances automatic encoding of melodic contour and interval structure. *J Cogn Neurosci*. 16:1010–1021.
- Grush R. 2004. The emulation theory of representation: motor control, imagery, and perception. *Behav Brain Sci*. 27:377–396; discussion 396–442.
- Hall DA, Haggard MP, Akeroyd MA, Palmer AR, Summerfield AQ, Elliott MR, Gurney EM, Bowtell RW. 1999. “Sparse” temporal sampling in auditory fMRI. *Hum Brain Mapp*. 7:213–223.
- Hanakawa T, Dimyan MA, Hallett M. 2008. Motor planning, imagery, and execution in the distributed motor network: a time-course study with functional MRI. *Cereb Cortex*. 18:2775–2788.
- Herholz SC, Halpern AR, Zatorre RJ. 2012. Neuronal correlates of perception, imagery, and memory for familiar tunes. *J Cogn Neurosci*. 24:1382–1397.
- Herholz SC, Lappe C, Knief A, Pantev C. 2008. Neural basis of music imagery and the effect of musical expertise. *Eur J Neurosci*. 28:2352–2360.

- Herholz SC, Zatorre RJ. 2012. Musical training as a framework for brain plasticity: behavior, function, and structure. *Neuron*. 76:486–502.
- Hickok G, Poeppel D. 2007. The cortical organization of speech processing. *Nat Rev Neurosci*. 8:393–402.
- Hoshi E, Tanji J. 2006. Differential involvement of neurons in the dorsal and ventral premotor cortex during processing of visual signals for action planning. *J Neurophysiol*. 95:3596–3616.
- Hubert V, Beaunieux H, Chetelat G, Platel H, Landeau B, Danion JM, Viader F, Desgranges B. 2007. The dynamic network subserving the three phases of cognitive procedural learning. *Hum Brain Mapp*. 28:1415–1429.
- Jäncke L. 2009. The plastic human brain. *Restor Neurol Neurosci*. 27:521–538.
- Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. 2012. *Fsl. Neuroimage*. 62:782–790.
- Johnsrude IS, Penhune VB, Zatorre RJ. 2000. Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain*. 123(Pt 1):155–163.
- Kohler E, Keysers C, Umiltà MA, Fogassi L, Gallese V, Rizzolatti G. 2002. Hearing sounds, understanding actions: action representation in mirror neurons. *Science*. 297:846–848.
- Krumbholz K, Patterson RD, Seither-Preisler A, Lammertmann C, Lutkenhoner B. 2003. Neuromagnetic evidence for a pitch processing center in Heschl's gyrus. *Cereb Cortex*. 13:765–772.
- Lahav A, Saltzman E, Schlaug G. 2007. Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *J Neurosci*. 27:308–314.
- Lappe C, Herholz SC, Trainor LJ, Pantev C. 2008. Cortical plasticity induced by short-term unimodal and multimodal musical training. *J Neurosci*. 28:9632–9639.
- May A. 2011. Experience-dependent structural plasticity in the adult human brain. *Trends Cogn Sci*. 15:475–482.
- Münste TF, Altenmüller E, Jäncke L. 2002. The musician's brain as a model of neuroplasticity. *Nat Rev Neurosci*. 3:473–478.
- Nedelko V, Hassa T, Hamzei F, Schoenfeld MA, Dettmers C. 2012. Action imagery combined with action observation activates more corticomotor regions than action observation alone. *J Neurol Phys Ther*. 36:182–188.
- Paraskevopoulos E, Kuchenbuch A, Herholz SC, Pantev C. 2012. Evidence for training-induced plasticity in multisensory brain structures: an MEG study. *PLoS One*. 7:e36534.
- Patterson RD, Uppenkamp S, Johnsrude IS, Griffiths TD. 2002. The processing of temporal pitch and melody information in auditory cortex. *Neuron*. 36:767–776.
- Penagos H, Melcher JR, Oxenham AJ. 2004. A neural representation of pitch salience in nonprimary human auditory cortex revealed with functional magnetic resonance imaging. *J Neurosci*. 24:6810–6815.
- Penhune VB, Steele CJ. 2012. Parallel contributions of cerebellar, striatal and M1 mechanisms to motor sequence learning. *Behav Brain Res*. 226:579–591.
- Petrides M. 1985. Deficits in non-spatial conditional associative learning after periarculate lesions in the monkey. *Behav Brain Res*. 16:95–101.
- Pizzamiglio L, Aprile T, Spitoni G, Pitzalis S, Bates E, D'Amico S, Di Russo F. 2005. Separate neural systems for processing action- or non-action-related sounds. *Neuroimage*. 24:852–861.
- Rauschecker JP, Scott SK. 2009. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat Neurosci*. 12:718–724.
- Schneider P, Scherg M, Dosch HG, Specht HJ, Gutschalk A, Rupp A. 2002. Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nat Neurosci*. 5:688–694.
- Schulze K, Zysset S, Mueller K, Friederici AD, Koelsch S. 2011. Neuroarchitecture of verbal and tonal working memory in nonmusicians and musicians. *Hum Brain Mapp*. 32:771–783.
- Shahin AJ, Bosnyak DJ, Trainor LJ, Roberts LE. 2003. Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. *J Neurosci*. 23:5545–5552.
- Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE. 2004. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*. 23: S208–S219.
- Stewart L, Henson R, Kampe K, Walsh V, Turner R, Frith U. 2003. Brain changes after learning to read and play music. *Neuroimage*. 20:71–83.
- Thomas A, Marrett S, Saad Z, Ruff D, Martin A, Bandettini P. 2009. Functional but not structural changes associated with learning: an exploration of longitudinal voxel-based morphometry (VBM). *NeuroImage*. 48:117–125.
- Warren JE, Wise RJ, Warren JD. 2005. Sounds do-able: auditory-motor transformations and the posterior temporal plane. *Trends Neurosci*. 28:636–643.
- Watanabe T, Yagishita S, Kikyo H. 2008. Memory of music: roles of right hippocampus and left inferior frontal gyrus. *Neuroimage*. 39:483–491.
- Wong PC, Warrier CM, Penhune VB, Roy AK, Sadeh A, Parrish TB, Zatorre RJ. 2008. Volume of left Heschl's Gyrus and linguistic pitch learning. *Cereb Cortex*. 18:828–836.
- Woolrich M. 2008. Robust group analysis using outlier inference. *Neuroimage*. 41:286–301.
- Woolrich MW, Behrens TE, Beckmann CF, Jenkinson M, Smith SM. 2004. Multilevel linear modelling for fMRI group analysis using Bayesian inference. *Neuroimage*. 21:1732–1747.
- Worsley K. 2001. Statistical analysis of activation images. *Functional MRI: an Introduction to Methods*. 14:251–270.
- Zatorre RJ. 1988. Pitch perception of complex tones and human temporal lobe function. *J Acoust Soc Am*. 84:566.
- Zatorre RJ. 2013. Predispositions and plasticity in music and speech learning: neural correlates and implications. *Science*. 342:585–589.
- Zatorre RJ, Belin P. 2001. Spectral and temporal processing in human auditory cortex. *Cereb Cortex*. 11:946–953.
- Zatorre RJ, Chen JL, Penhune VB. 2007. When the brain plays music: auditory-motor interactions in music perception and production. *Nat Rev Neurosci*. 8:547–558.
- Zatorre RJ, Delhommeau K, Zarate JM. 2012. Modulation of auditory cortex response to pitch variation following training with microtonal melodies. *Front Psychol*. 3:544.
- Zatorre RJ, Fields RD, Johansen-Berg H. 2012. Plasticity in gray and white: neuroimaging changes in brain structure during learning. *Nat Neurosci*. 15:528–536.