

# Training-mediated leftward asymmetries during music processing: A cross-sectional and longitudinal fMRI analysis

Robert J. Ellis<sup>a</sup>, Bente Bruijn<sup>b</sup>, Andrea C. Norton<sup>a</sup>, Ellen Winner<sup>c</sup>, Gottfried Schlaug<sup>a,\*</sup>

<sup>a</sup> Department of Neurology, Beth Israel Deaconess Medical Center and Harvard Medical School, 330 Brookline Ave, Palmer 127, Boston, MA 02215, USA

<sup>b</sup> Medical Faculty AMC-UvA, University of Amsterdam, Meibergdreef 9, 1105 AZ, Amsterdam, The Netherlands

<sup>c</sup> Department of Psychology, Boston College, McGuinn Hall, 140 Commonwealth Ave., Chestnut Hill, MA 02467, USA

## ARTICLE INFO

### Article history:

Accepted 20 February 2013

Available online 5 March 2013

### Keywords:

Auditory discrimination  
Hemispheric lateralization  
Musical training  
Working memory

## ABSTRACT

Practicing a musical instrument has a profound impact on the structure and function of the human brain. The present fMRI study explored how relative hemispheric asymmetries in task-related activity during music processing (*same/different* discrimination) are shaped by musical training (quantified as cumulative hours of instrument practice), using both a large ( $N = 84$ ) cross-sectional data set of children and adults, and a smaller ( $N = 20$ ) two time-point longitudinal data set of children tracked over 3 to 5 years. The cross-sectional analysis revealed a significant leftward asymmetry in task-related activation, with peaks in Heschl's gyrus and supramarginal gyrus (SMG). The SMG peak was further characterized by a leftward asymmetry in the partial correlation strength with subjects' cumulative hours of practice, controlling for subjects' age and task performance. This SMG peak was found to exhibit a similar pattern of response in the longitudinal data set (in this case, with subjects' cumulative hours of practice over the course of the study), controlling for age, scan interval, and amount of instrument practice prior to the first scan. This study presents novel insights into the ways musical instrument training shapes task-related asymmetries in neural activity during music processing.

© 2013 Elsevier Inc. All rights reserved.

## Introduction

Learning a musical instrument places many demands on the human brain: fine-grained coordination within and between articulatory systems; real-time adjustments in response to auditory, visual, and somatosensory cues; and (when required) the ability to perform these actions in synchrony with other individuals. These perceptual, cognitive, and motor demands have translated into a bounty for the field of neuroscience (for recent reviews, see Jäncke, 2009; Penhune, 2011; Wan and Schlaug, 2010; Zatorre et al., 2007). A particular benefit with respect to experimental design and analysis is that musical training can be easily parameterized, isolating a specific feature such as the age at which training commenced (e.g., Amunts et al., 1997; Pantev et al., 1998), the number of years spent training (e.g., Musacchia et al., 2007), the cumulative hours of practice (e.g., Bengtsson et al., 2005; Kleber et al., 2010), or some combination of these features (e.g., Elmer et al., 2012; Foster and Zatorre, 2010; Hutchinson et al., 2003; Ohnishi et al., 2001; Sluming et al., 2002). Treating musical training as a continuum rather than a category (i.e., “musicians vs. nonmusicians”) prevents well-known losses in statistical power (e.g., MacCallum et al.,

2002). More importantly, it affords some insight into *which* aspects of training (e.g., age of commencement versus duration or intensity of practice) relate to specific differences in brain structure or brain function—something not possible with categorical designs.

In our own previous functional magnetic resonance imaging (fMRI) investigation (Ellis et al., 2012) using a musical phrase discrimination task, multiple regression analysis revealed a positive partial correlation between subjects' cumulative hours of musical training (controlling for age and task performance) and activity in left posterior superior temporal gyrus (pSTG)/planum temporale (PT). The PT plays a role in sequencing spectrotemporal patterns and comparing them to stored templates (Griffiths and Warren, 2002), facilitating an auditory input/motor output coordinate transformation (Warren et al., 2005)—actions in which trained musicians might preferentially engage when discriminating melodies or rhythms. The asymmetric nature of this correlation (i.e., significant on the left, non-significant on the right) is consistent with previous reports revealing leftward asymmetries in PT surface area in musicians with absolute pitch (Keenan et al., 2001; Schlaug et al., 1995b), and a negative correlation between the age of commencement of training and activity in left PT during passive listening to music (Ohnishi et al., 2001). It is also in agreement with studies reporting *relative* leftward asymmetries (i.e., based on a direct comparison of effect magnitude on the left and right, rather than an isolated effect on the left) that are stronger in (categorically defined)

\* Corresponding author. Fax: +1 617 632 8920.

E-mail address: [gschlaug@bidmc.harvard.edu](mailto:gschlaug@bidmc.harvard.edu) (G. Schlaug).

musicians compared to nonmusicians (e.g., Bever and Chiarello, 1974; Elmer et al., 2012; Herholz et al., 2011; Tervaniemi et al., 2011).

Our previous left-hemisphere finding, however, must be interpreted with (at least) three caveats. First, the degree to which any effect (either a main effect or a correlation) appears asymmetric or lateralized in a statistical parametric mapping (SPM) analysis depends strongly on how that map was thresholded (e.g., Abbott et al., 2010; Jansen et al., 2006). Second, a significant partial correlation (i.e., a regression slope  $t$ -test) is statistically independent of any “main effect” (i.e., a significant intercept  $t$ -test); only the latter indicates that a clear majority of subjects (i.e., assuming typical voxel-level thresholds at or below  $p < .005$ ) actually showed the same pattern of activity (e.g., positive values during *Task > Rest*). Third, musical training was assessed in a cross-sectional design rather than a longitudinal design; only the latter is able to accurately characterize training-induced neuroplasticity (e.g., Jäncke, 2009; Peretz and Zatorre, 2005).

The present study was constructed to address these issues. To address the first two caveats, the cross-sectional data in Ellis et al. (2012) were reanalyzed to identify regions of interest (ROIs) which satisfied a series of conditionals, illustrated in Fig. 1. First, a standard whole-brain SPM  $t$ -test was performed to identify regions showing a significant overall task-related response (Conditional 1; e.g., “ $t_L > 0$ ” for left hemisphere voxels). Second, a whole-brain “lateralized SPM”  $t$ -test (e.g., Liégeois et al., 2002; Stevens et al., 2005) was performed to identify regions showing a significant overall hemispheric asymmetry in task-related response (Conditional 2; e.g., “ $t_L > t_R$ ”). Third, regions satisfying Conditionals 1 and 2 were subjected to offline robust regressions (for a discussion, see Wager et al., 2005) to quantify partial correlations with musical training (controlling for age and scanner task performance), separately in each hemisphere. Fourth, the difference in correlation strength between the hemispheres was evaluated using a Z-test for dependent correlations (Meng et al., 1992) to identify a significant difference (i.e., hemispheric asymmetry) in correlation strength (Conditional 3; e.g., “ $r_L > r_R$ ”). Thus, regions which satisfied all three conditionals exhibited both an overall asymmetry in task-related response, as well as an asymmetry in the strength of the correlation between that response and inter-subject differences in musical training (i.e., Fig. 1d).

To address the third caveat, the cross-sectional analysis was paired with a longitudinal analysis of the development of training-mediated asymmetries, using a 20-subject data set (detailed in the Materials

and methods section) collected as part of a larger investigation of neural, behavioral, and cognitive changes associated with long-term musical training in children (e.g., Norton et al., 2005; Schlaug et al., 2005; Schlaug et al., 2009). Previous longitudinal studies in children (as well as adults: e.g., Bangert and Altenmüller, 2003; Jabusch et al., 2009) have explored how musical training affects performance on cognitive (e.g., Ho et al., 2003; Schellenberg, 2004) and motor (Forgeard et al., 2008) tasks, electroencephalographic (e.g., Besson et al., 2011; Moreno et al., 2009) and magnetoencephalographic (Fujioka et al., 2006) signals, and morphometric changes in primary auditory and motor areas (Hyde et al., 2009). The present study is the first reported fMRI analysis of longitudinal changes associated with musical training in children.

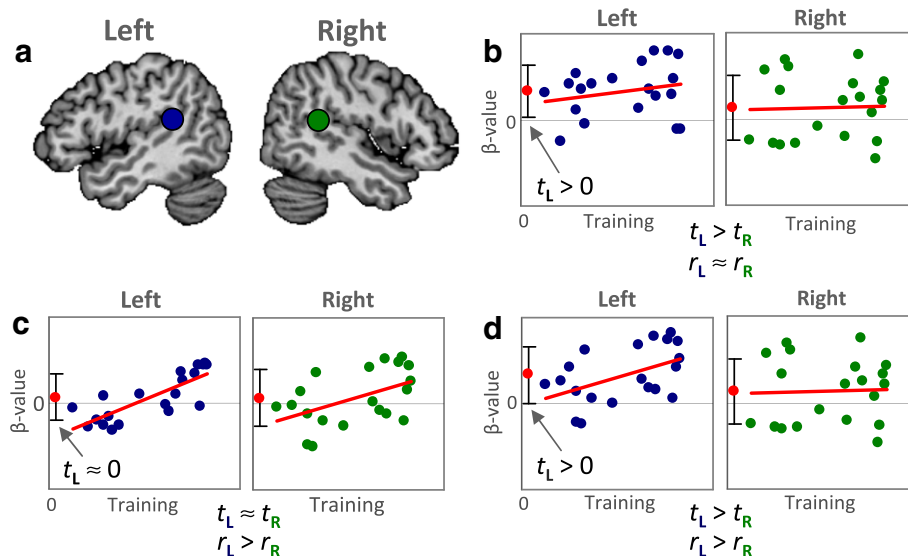
Furthermore, the full combination of voxelwise laterality assessments of fMRI activity and parametric analyses of musical training assessed both cross-sectionally and longitudinally is a unique one. A previous study (Herdener et al., 2010) examined cross-sectional and longitudinal effects of musical training (in adult subjects only) on hippocampal activity during passive listening to rapid tone sequences with temporal deviants. Our study explores training-mediated asymmetries in fMRI activity (in both children and adults) during an *active* musical discrimination task.

## Materials and methods

### Subjects

A cross-sectional and a longitudinal data set were used in the present analysis. The cross-sectional data set comprised the same 84 subjects analyzed in Ellis et al. (2012): 28 adults (aged 21–33) and 28 children (aged 9–11) who participated in the study's cross-sectional arm, and 28 children (aged 5–7 at first scan) selected from the study's longitudinal arm. Table 1a provides cross-sectional subject demographics. Within each age group, half ( $N = 14$ ) of the subjects had received no musical training at the time of scan.

The longitudinal data set comprised 20 children who participated in a multi-year investigation of neural, behavioral, and cognitive changes associated with training on a musical instrument. (For more details about this data set, see Norton et al., 2005; Schlaug et al., 2005; Schlaug et al., 2009.) All longitudinal participants were 5–7 years old at the time of enrollment in the study, and all received 30–40 min per week of general music in school. A subset of children who were about to begin (or had recently begun) weekly private



**Fig. 1.** Illustration of the conditionals used to identify a bilateral region of interest (visualized in panel a) in the large cross-sectional data set. Conditional 1 requires a significant overall task-related effect, symbolized by as “ $t_L > 0$ ” for a significant left-hemisphere effect (present in panels b and d). Conditional 2 requires a significant asymmetry in overall task-related effect, symbolized by “ $t_L > t_R$ ” for a significant leftward asymmetry (present in panels b and d). Conditional 3 requires a significant difference (asymmetry) in the correlation strength with a given regressor (e.g., for hours of musical training), symbolized as “ $r_L > r_R$ ” (present in panels c and d). Only panel d satisfies all three conditionals.

**Table 1a**

Demographics for the 84 subjects in the cross-sectional data set, grouped as a function of age and training. T: trained; U: untrained.

Group	Males/females	Mean (SD)/range			
		Age at commencement	Age at scan	Years of practice	Hours of practice
5-to-7 U	7/7	–	6.23 (0.56)/5.10–7.11	0	0
5-to-7 T	5/9	6.03 (0.77)/4.84–6.91	6.46 (0.80)/5.04–7.38	0.38 (0.17)/0.16–0.82	15.57 (8.86)/4.94–35.91
9-to-11 U	8/6	–	10.07 (0.67)/9.10–11.16	0	0
9-to-11 T	5/9	5.74 (1.35)/4.01–8.46	10.28 (0.73)/9.08–11.12	4.61 (1.61)/1.24–6.63	1535 (1144)/268–3792
Adult U	6/8	–	27.17 (3.50)/20.96–33.00	0	0
Adult T	8/6	5.21 (1.05)/4.00–8.00	25.87 (2.85)/21.51–31.33	19.07 (4.51)/10.00–26.00	10,888 (5053)/4473–20,849

lessons with a piano or stringed instrument teacher (and who planned to practice regularly at home) were assigned to the *Instrumental* group. Children whose families had not opted for private study or practice were assigned to the *Control* group. The study had a rolling enrollment; two years were required to enroll all children, and a number of the children who had initially been enrolled as control subjects began to study an instrument. It was thus necessary to combine all children into a single analysis group, and use a set of regressors to model subjects' demographics (age, total hours of instrumental practice) and task performance, as discussed in detail below.

In the present analysis, data from two imaging time points per child (referred to here as “Baseline” and “Final”) was used, with specific demographic and timing details provided in Table 1b. Of the 20 subjects, five had received no musical training at either Baseline or Final; five had begun training after their Baseline scan; and ten had some amount of training at Baseline and were continuing to train at the time of the Final scan.

Hours of practice was quantified as the cumulative number of hours of instrument instruction, ensemble practice/performance (if applicable), and individual practice time since commencement of musical training (cf. Bengtsson et al., 2005; Hutchinson et al., 2003; Kleber et al., 2010). For subjects in the cross-sectional study, this was calculated from information provided (retrospectively) on questionnaires; for subjects in the longitudinal study, this was calculated based on information from parent questionnaires (to tally any instruction and practice time prior to the Baseline scan) and from weekly practice records kept in journals and reported on a monthly basis (to quantify training intensity between the Baseline and Final scans). Finally, handedness was assessed using measures adapted from Annett (1970) by Norton et al. (2005). All subjects in the cross-sectional study and all children in the longitudinal study (tested at both time points) were classified as consistently right handed.

### Stimuli and tasks

The scanner task comprised a *same/different* melodic discrimination (MD) or rhythmic discrimination (RD) judgment of pairs of five-note musical phrases via a button press with the index finger of the left (“same”) or right (“different”) hand, and is described in greater detail in Overy et al. (2004). All MD stimuli had an isochronous rhythm (i.e., constant eighth notes); all RD stimuli had a non-isochronous rhythm. All stimulus pairs can be accurately notated using a 4/4 metrical grid, as illustrated in Fig. 2. A single run consisted of 12 trials: eight phrase pairs (either MD or RD) and four silent (S) trials, during which subjects heard no musical stimulus, but made a bimanual button press after a noise burst cue.

Within each run, 3 *same* and 5 *different* phrase pairs were presented; all subjects successfully completed four runs (2 MD and 2 RD, in alternation). Subjects were familiarized with the discrimination task and trial procedure during a behavioral testing session approximately one week prior to scanning. Scanner performance was quantified via the sensitivity index  $d'$  (e.g., Macmillan and Creelman, 2005, Ch. 2):  $Z(\text{hit rate}) - Z(\text{false alarm rate})$ .

### Regressor specification

The cross-sectional model utilized three regressors, calculated at the time of the scan: age (“ $Age_S$ ”), hours of practice (“ $HrsPract_S$ ”), and scanner task  $d'$  (“ $Performance_S$ ”). The longitudinal regression model used a set of five regressors. Three quantified change from Baseline to Final: in age (“ $Age_{F-B}$ ”), total hours of instrument practice (“ $HrsPract_{F-B}$ ”), and  $d'$  (“ $Performance_{F-B}$ ”). The other two quantified task-independent differences at Baseline: in age (“ $Age_B$ ”), and hours of practice (“ $HrsPract_B$ ”). Including these latter regressors enabled inter-subject differences that were (statistically) constant at both imaging time points to be isolated from inter-subject differences that changed between the time points. All  $Age$  and  $HrsPract$  regressors (in both models) were subjected to a natural log transform ( $\ln(\text{value} + 1)$ ), which yielded higher Shapiro–Wilk  $W$  values (i.e., less departure from normality) than when untransformed. Following this, the variance inflation factor (VIF) for each regressor in both models was low (all VIFs < 2.4 for the cross-sectional model, and all VIFs < 2.2 for the longitudinal model), indicating a lack of multicollinearity (e.g., Kutner et al., 2004).

We note that it was not possible to include the age of commencement of training in the regression models for statistical reasons. Although “Hours of practice = 0” is meaningful, “Age of commencement = 0” is not (cf. Poldrack et al., 2011, Ch. 6.2), precluding those subjects from any regression model in which age of commencement was a factor. Additionally, it was found that multicollinearity was present when age of commencement was included in the regression model on the 42 subjects with training, due to a strong correlation between  $Age_S$  and  $HrsPract_S$  ( $r = .898$ ) that was less pronounced when all 84 subjects were available for use ( $r = .328$ ).

### Image acquisition

Functional images were acquired via a sparse sampling design (e.g., Belin et al., 1999; Gaab et al., 2003) on a 3 T General Electric magnetic resonance imaging scanner using a gradient-echo EPI-sequence with an echo time of 25 ms and a  $64 \times 64$  mm matrix. Using a mid-sagittal scout image, 26 slices were acquired over 1.75 s with a

**Table 1b**

Demographics for the 20 subjects in the longitudinal data set, grouped to illustrate differences in the amount of training at Baseline and Final.

Group	Males/females	Mean (SD)/range				
		Age at commencement	Age at baseline	Age at final	Hours of practice at baseline	Hours of practice at final
A	1/4	–	5.94 (.46)/5.42–6.52	9.58 (.89)/8.32–10.57	0	0
B	1/4	8.43 (1.69)/5.89–9.96	6.34 (.84)/5.53–7.63	10.59 (.99)/9.35–11.92	0	251.26 (233.90)/40.57–622.50
C	3/7	6.02 (.85)/4.84–7.48	6.47 (.87)/9.32–11.42	10.24 (.78)/9.32–11.42	36.12 (20.65)/6.32–70.67	348.35 (147.53)/82.60–526.11

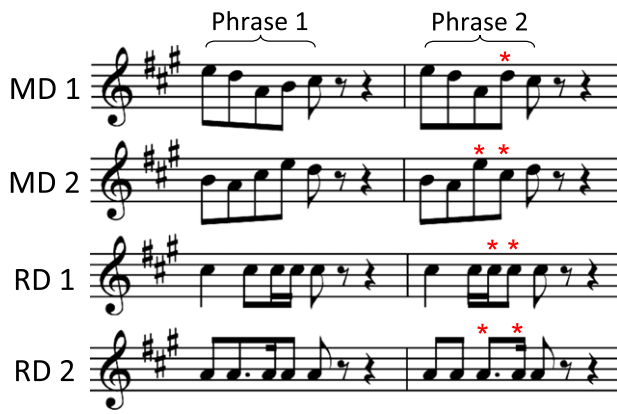


Fig. 2. Sample stimuli for the melodic and rhythmic discrimination task. Asterisks indicate a change in Phrase 2 relative to Phrase 1.

voxel size of  $3.8 \times 3.8 \times 4$  mm. Scanning repetition time (TR) was kept constant at 15 s; stimuli were jittered between three time-points such that the onset of the first axial slice occurred 1.25, 2.25, or 3.25 s after the end of the second phrase in each pair. The data from these jitter points were combined during statistical analysis to allow for individual differences in hemodynamic response time across brain regions.

#### Image preprocessing and first-level analysis

Data were preprocessed using a standard pipeline (realignment, movement correction, normalization to the MNI EPI template, smoothing with an isotropic 8-mm FWHM kernel, and resampling to 2-mm cubic voxels) using the SPM5 software suite ([www.fil.ion.ucl.ac.uk/spm/](http://www.fil.ion.ucl.ac.uk/spm/)). With respect to movement correction, across all individual functional runs and all subjects, the standard deviation (within each run) was less than 2.5 mm across translation parameters, and less than  $0.1^\circ$  across rotation parameters.

First-level analysis was performed using a finite-impulse response basis function (window length = 1 s, order = 1). Global normalization (scaling) was applied, and serial correlations in the fMRI time series were modeled using a first-order autoregressive algorithm. Low frequency drifts were removed using a temporal high-pass filter with a cutoff of 200 s; no low-pass filter was applied. A box-car function was applied with an epoch length of 1 to the fMRI time series (12 acquisitions within each run: 8 MD or RD, 4 S), and no temporal derivatives were applied.

#### First-level contrasts

Three first-level contrasts were created for each subject: one for melodic discrimination ( $MD > S$ ), one for rhythmic discrimination ( $RD > S$ ), and one for “average” discrimination across all MD and RD runs ( $AD > S$ ). Because we did not manipulate fine-grained melodic and rhythmic properties in our stimuli (e.g., Schönwiesner et al., 2005; Zatorre and Belin, 2001) we anticipated (and found) largely similar patterns of activity in  $MD > S$  and  $RD > S$  (Ellis et al., 2012). For this reason, initial SPM analyses were performed on  $AD > S$  contrasts to take advantage of the increased power (i.e., more runs). Offline analyses (regression and ANOVA) were then performed (on the large cross-sectional data set only) to explore potential differences between  $MD > S$  and  $RD > S$ .

#### Renormalization and lateralized contrast specification

In a typical lateralized SPM pipeline (e.g., Liégeois et al., 2002; Stevens et al., 2005; Swanson et al., 2011), realigned scans are normalized to a symmetrical template (i.e., the average of a standard template and an x-flipped version of that template) prior to first-level analysis.

Our own data, however, had already been normalized to the standard (nonsymmetrical) EPI template prior to formulating the current experimental question. Thus, a symmetrical renormalization step (cf. Büchel et al., 2004; Luders et al., 2004) was performed by: (1) normalizing each subject's realigned mean EPI image to the standard template; (2) estimating the parameters necessary to renormalize these images to the symmetrical EPI template, and (3) applying those unique parameters to that subject's  $AD > S$ ,  $MD > S$ , and  $RD > S$  contrasts.

After renormalization, a lateralized version of each contrast was created by subtracting an x-flipped version of the image from the unflipped (original) image. A positive value for a given voxel indicates relatively greater activity at that voxel versus its homotopic (contralateral) voxel. Three unique lateralized contrasts were created for each of the 84 subjects in the cross-sectional analysis, and for the 20 subjects at each time point in the longitudinal analysis.

#### Second-level SPM analyses, visualization, and cluster labeling

As outlined in the Introduction, SPM second-level *t*-tests were performed on unlateralized (Conditional 1) and lateralized (Conditional 2)  $AD > S$  contrasts. All SPMs were thresholded at voxel-level  $p < .001$  and FWE-corrected cluster-level  $p < .05$ , per random field theory assumptions (actual extent thresholds were queried using *CorrClusTh*; Nichols, 2008).

Anatomical localization of resultant clusters was performed using the Anatomy toolbox (version 1.8; Eickhoff et al., 2005), which provides gray matter cytoarchitectonic probabilities (where available) and macroanatomic labels borrowed from the Automated Anatomical Labeling parcellation of the MNI single-subject template (Tzourio-Mazoyer et al., 2002). Macro- or micro-anatomical labeling of clusters is more challenging with lateralized SPM clusters, as well-known interhemispheric differences exist (e.g., Lyttelton et al., 2009; Toga and Thompson, 2003). For this reason, each lateralized cluster was visualized and parcellated in both hemispheres (i.e., at its original location and mirrored about the x-axis). Final visualizations were prepared using MRIcron ([www.mccauslandcenter.sc.edu/mricron/mricron/](http://www.mccauslandcenter.sc.edu/mricron/mricron/)).

#### Offline robust regressions and correlation comparisons

A standard multiple regression analysis (i.e., ordinary least squares, OLS) is sensitive to the presence of outliers that can inflate or diminish resultant correlations (e.g., Yarkoni, 2009). To combat this, a robust regression approach (i.e., iteratively reweighted least squares, IRLS; e.g., Wager et al., 2005) was performed (using *robustfit* in MATLAB, with all regressors mean centered). That is, IRLS was used to calculate the strength of regression slopes for all factors in the cross-sectional and longitudinal regression models, for regions of interest satisfying Conditionals 1 and 2 identified from the whole-brain SPM analyses. A Z-test (Meng et al., 1992) was then performed to determine if the left- and right-hemisphere regression slopes were significantly different (Conditional 3). This offline analysis is not circular, as the strength of a correlation is statistically independent from a one-sample *t*-test on those same values (i.e., the SPM results). (Note: although Fig. 1 reflects the expected, leftward asymmetry in response, both *Left > Right* and *Right > Left* effects are possible using this three-conditional approach.)

## Results

#### Scanner task performance

Scanner task performance was also analyzed using robust regression. In the cross-sectional sample, a two-factor regression was performed with *Performance<sub>s</sub>* as the dependent variable. A significant main effect (i.e., intercept *t*-test) was present, indicating greater-than-chance *d'* sensitivity across all subjects (i.e.,  $d' > 0$ ;  $t_{81} = 25.09$ ,  $p < .0001$ ). Significant partial correlations with *Performance<sub>s</sub>* (i.e., regression slope *t*-tests) were



present for both  $Age_S$  ( $r_{81} = .651, p < .0001$ ) and  $HrsPract_S$  ( $r_{81} = .407, p = .0001$ ). In the longitudinal sample, a four-factor regression was performed with  $Performance_{F-B}$  as the dependent variable. A significant increase in  $d'$  sensitivity (from Baseline to Final) was present (intercept  $t_{15} = 6.73, p < .0001$ ), and a near-significant partial correlation was found for  $Age_{F-B}$  ( $r_{15} = .468, p = .058$ ). Partial correlations for  $HrsPract_{F-B}$  ( $r_{15} = -.061, p = .816$ ),  $Age_B$  ( $r_{15} = -.195, p = .453$ ), and  $HrsPract_B$  ( $r_{15} = -.080, p = .760$ ) were not significant. (Note: because of the large differences in  $N$  between the two studies, direct comparisons of  $t$ - or  $r$ -values are not fruitful;  $p$ -values reflect the actual statistical significance of a given effect.)

A set of post-hoc analyses examined potential differences in  $d'$  as a function of stimulus type (MD vs. RD). In the cross-sectional sample, a paired  $t$ -test comparing  $Performance_S$  during MD and RD was not significant ( $t_{83} = 1.13, p = .259$ ); that is, the intercepts from separate two-factor robust regressions on MD and RD  $d'$  were similar (MD:  $t_{81} = 23.14$ ; RD:  $t_{81} = 21.74$ ). The magnitude of the partial correlation for  $Age_S$  was similar during MD ( $r_{81} = .595, p < .001$ ) and RD ( $r_{81} = .573, p < .001$ ). The partial correlation for  $HrsPract_S$  was somewhat larger during MD ( $r_{81} = .439, p < .001$ ) than RD ( $r_{81} = .324, p = .003$ ); however, these two were not significantly different per the Meng et al. (1992)  $Z$ -test ( $Z = 1.45, p = .147$ ).

In the longitudinal sample, a paired  $t$ -test comparing  $Performance_{F-B}$  during MD and RD was not significant ( $t_{19} = -.89, p = .381$ ); similar to the cross-sectional design, the intercept of the four-factor robust regression (reflecting an overall increase in  $d'$  from Baseline to Final) was similar during MD ( $t_{15} = 4.94, p < .001$ ) and RD ( $t_{15} = 4.19, p < .001$ ). The partial correlation with  $HrsPract_{F-B}$  (that was near-significant when MD and RD was pooled) was somewhat larger during MD ( $r_{15} = .417, p = .096$ ) than RD ( $r_{15} = .184, p = .271$ ), but this difference was not significant ( $Z = .725, p = .468$ ).

#### Cross-sectional SPM results

A single large cluster (959 voxels at  $p_{\text{voxel}} < .001$ ) emerged during the conjunction of the unlateralized  $AD > S$  and lateralized  $AD > S$  SPMs (i.e., Conditionals 1 and 2), and is visualized in Fig. 3a. This cluster had two spatially distinct peaks in the lateralized SPM (Peak 1:  $t_{83} = 5.48$  at  $\{-52, -42, 22\}$ ; Peak 2:  $t_{83} = 5.17$  at  $\{-48, -24, 6\}$ ) that remained conjoined even at a more stringent threshold (486 voxels at  $p_{\text{voxel}} < .0001$ ). To facilitate anatomical inference without further “jittering” the threshold (cf. Poldrack, 2007), a sphere (radius = 8 mm; 257 voxels) was drawn at each peak (Fig. 3b), mirrored about the  $x$ -axis (to enable localization in both hemispheres), and parcellated using the Anatomy toolbox (Table 2). In both hemispheres, the Peak 1 sphere overlapped most extensively (63.5% on the left and 66.0% on the right) with cytoarchitectonic subdivisions of ventral supramarginal gyrus (SMG), and the Peak 2 sphere overlapped most extensively (60.0% on the left and 50.7% on the right) with subdivisions of Heschl's gyrus (HG; primary auditory cortex). This result motivated the labels “SMG” and “HG” for the Peak 1 and Peak 2 clusters, respectively.

#### Cross-sectional: SMG and HG ROI analyses

Offline robust regressions were performed to quantify the strength of partial correlations ( $Age_S$ ,  $HrsPract_S$ ,  $Performance_S$ ) in each hemisphere and each ROI. Separate regressions were performed on  $\beta$ -values (extracted from each subject's unlateralized  $AD > S$  contrast) from the left and right versions of the SMG and HG ROIs. Table 3 summarizes these results, and the subsequent  $Z$ -test to evaluate the significance of hemispheric differences (asymmetries) in correlation strength (cf. Fig. 1). As shown in Fig. 3c, the most interesting result from this analysis was found in the SMG (but not the HG) ROI: a significant hemispheric asymmetry ( $Z = 2.99, p = .003$ , two-tailed) in the strength of the left-vs.-right partial correlation of  $\beta$ -values with  $HrsPract_S$ . (Partial

correlations plot  $\beta$ -value residuals on the ordinate against  $HrsPract_S$  residuals on the abscissa, with higher values indicating greater  $HrsPract_S$ ).

To further characterize the actual task-related response (i.e., raw  $\beta$ -values rather than residualized  $\beta$ -values) in left SMG, a 3 ( $Age$ : 5-to-7 vs. 9-to-11 vs. Adult)  $\times$  2 ( $HrsPract$ : Untrained vs. Trained) analysis of variance (ANOVA) was performed, and is visualized in Fig. 3d. A post-hoc,  $Age_{\text{linear}} \times Hemisphere$  interaction achieved near significance ( $F_{1,78} = 3.39, p = .068, \eta^2_p = .042$ ), suggesting an even more nuanced response in left SMG: specifically, decreasing  $\beta$ -values as subjects without musical training matured, and increasing  $\beta$ -values as subjects with musical training matured. That is, mean  $\beta$ -values in left SMG were maximally different between the Untrained Adult and Trained Adult groups ( $F_{1,78} = 5.48, p = .022, \eta^2_p = .066$ ). (Partial eta-squared,  $\eta^2_p$ , quantifies the proportion of variance due to this effect:  $SS_{\text{effect}} / (SS_{\text{effect}} + SS_{\text{error}})$ ; e.g., Keppel and Wickens, 2004.)

#### Longitudinal: SMG ROI analysis

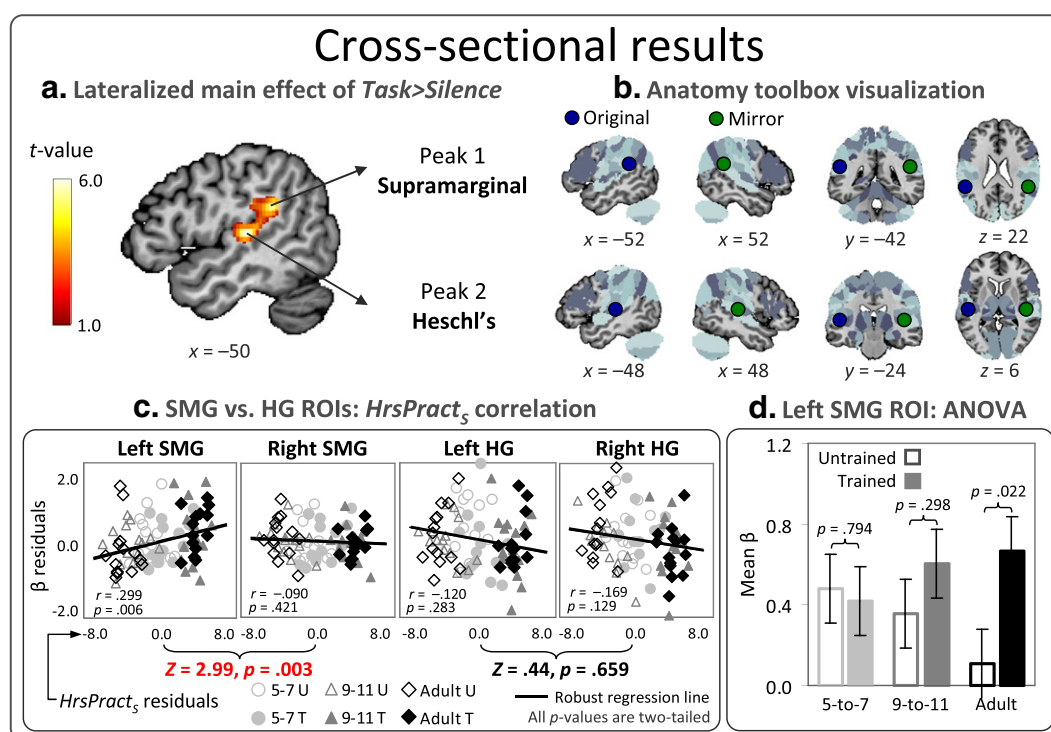
For sake of completeness, unlateralized and lateralized SPM  $t$ -tests (Conditionals 1 and 2) were performed on the set of 20 Final-minus-Baseline  $AD > S$  contrasts. No significant clusters were present at  $p_{\text{voxel}} < .001$  plus  $p_{\text{cluster}} < .05$  ( $k \geq 131$  voxels); or at  $p_{\text{voxel}} < .005$  plus  $p_{\text{cluster}} < .10$  ( $k \geq 332$  voxels), possibly due to fewer degrees of freedom ( $t_{19}$  vs.  $t_{81}$ ). As outlined in the Introduction, however, our primary use for the longitudinal data set was to test whether asymmetries in task-related activity associated with differences in musical training across subjects would also be associated with differences in musical training within subjects. The SMG ROI defined using the set of three conditionals (Fig. 1d) was used for this purpose.

Mean  $\beta$ -values from the left and right spherical SMG ROIs were extracted from each subject's (renormalized) unlateralized  $AD > S$  contrast at Baseline and Final. To characterize the overall task-related activity pattern, a 2 ( $Time$ : Baseline vs. Final)  $\times$  2 ( $Hemisphere$ : Left vs. Right) repeated-measures ANOVA was performed, with results shown in Fig. 4a. Consistent with the cross-sectional study, a significant asymmetry was present (i.e., a main effect of  $Hemisphere$ ;  $F_{1,19} = 11.24, p = .003$ ), that was larger at Final ( $F_{1,19} = 14.87, p = .001, \eta^2_p = .441$ ) than at Baseline ( $F_{1,19} = 7.88, p = .011, \eta^2_p = .293$ ). Although the  $Hemisphere \times Time$  interaction itself was not significant ( $F_{1,19} = 2.34, p = .142, \eta^2_p = .109$ ), the direction of the effect (i.e., larger asymmetry at Final relative to Baseline) is consistent with the cross-sectional result, and motivated further exploration using offline regression analyses.

Five-factor robust regressions were performed separately on left and right SMG ROI  $\beta$ -values. Fig. 4b visualizes partial correlations with each factor in the left and right hemisphere (cf. Fig. 1); supporting statistics are presented in Table 3. The  $Z$ -test was then performed on each pair of partial correlations to test for significant hemispheric asymmetries (while controlling for the remaining factors). Only the partial correlation with  $HrsPract_{F-B}$  was significantly asymmetric ( $Z = 2.07, p = .039$ , two-tailed). Similar to the cross-sectional  $HrsPract_S$  factor,  $HrsPract_{F-B}$  showed a positive correlation with left-hemisphere  $\beta$ -values, and a null correlation with right-hemisphere  $\beta$ -values.

#### Melodic vs. rhythmic discrimination

A final set of (planned) offline analyses examined whether training-mediated asymmetries in the SMG ROI was modulated by the type of discrimination subjects performed (melodic vs. rhythmic). To reduce the risk of a Type II error, only the large- $N$  cross-sectional data set was examined (i.e., no apparent differences in activation between MD and RD asymmetries in the 20-subject longitudinal data set may simply due to a lack of power). Mean  $\beta$ -values were extracted from the left and right hemisphere ROIs from each subject's unlateralized  $MD > S$  and  $RD > S$  contrast, and three-factor robust regressions were performed to quantify partial correlations with  $HrsPract_S$ .



**Fig. 3.** Results of the lateralized SPM analysis for the cross-sectional data set. (a) SPM visualization of the single significant cluster, reflecting a significant leftward asymmetry in activity during the discrimination task, with two spatially distinct peaks. (b) Anatomy toolbox visualization of the two peaks. A spherical region of interest (radius = 8 mm) was drawn at each original peak, mirrored in the right hemisphere, and parcellated using the Anatomy toolbox (cf. Table 2), yielding the labels supramarginal gyrus and Heschl's gyrus. (c) Partial correlation between cumulative hours of practice and mean  $\beta$ -values for the SMG and HG ROIs. (d) ANOVA results highlighting left-vs.-right differences in mean  $\beta$ -value as a function of age and training group (cf. Table 1a). Two-tailed  $p$ -values reflect planned contrasts (all  $F_{1,78}$ ).

To compare overall task-related activity, a 2 (Task: MD vs. RD)  $\times$  2 (Hemisphere: Left vs. Right) ANOVA was performed. A significant effect was found for Hemisphere ( $F_{1,83} = 29.30$ ,  $p < .0001$ ,  $\eta^2_p = .261$ ) but not Task ( $F_{1,83} = .11$ ,  $p = .744$ ) or Hemisphere  $\times$  Task ( $F_{1,83} < .01$ ). That is, the type of discrimination performed (melodic versus rhythmic) did not significantly impact activation asymmetries in the SMG ROI. (We consider a potential confound—that is, differences in the temporal complexity of MD and RD stimuli—in the Discussion.)

To compare hemispheric differences in partial correlation strength with *HrsPract<sub>s</sub>*, left-vs.-right Z-tests were performed separately for MD and RD. Additionally, a MD-vs.-RD Z-test was performed on left-hemisphere  $\beta$ -values, given that the influence of training was

stronger in the left hemisphere overall (cf. Fig. 3c) and thus might show a differential effect of stimulus type.

Table 4 summarizes the relevant statistics, including the Z-test results (all two-tailed). A significant asymmetry was present (positive correlation on the left; null correlation on the right) during both MD ( $p = .002$ ) and RD ( $p = .006$ ). The strength of the correlation with *HrsPract<sub>s</sub>* in the left SMG ROI was not significantly different during MD and RD ( $p = .625$ ).

## Discussion

The current study presents an in-depth examination of how training on a musical instrument shapes hemispheric asymmetries in fMRI activity when listeners discriminate pairs of melodies and rhythms. A large cross-sectional (between-subjects) analysis revealed a strongly left-lateralized activation with peaks in Heschl's gyrus (HG) and ventral supramarginal gyrus (SMG). Isolating the SMG peak revealed a significant difference (left > right) in the strength of the correlation between subjects' cumulative hours of practice and task-related activity (controlling for age and scanner task performance). Examining this relationship separately for melodic (MD) and rhythmic (RD) discrimination confirmed this leftward asymmetry during both tasks.

A more direct test of the role of musical training on task-related asymmetries in SMG was performed using a longitudinal (within-subjects) analysis comparing pairs of scans ("Baseline" and "Final") from children enrolled in a multi-year study. Similar to the cross-sectional result, SMG showed a significant asymmetry (left > right) in overall activity, as well as a significant difference (left > right) in the strength of the correlation with subjects' cumulative hours of practice between Baseline and Final (controlling for maturation effects, inter-subject differences in hours practiced prior to Baseline, and task performance). We distinguish these training-induced asymmetries from those observed when listeners

**Table 2**

Localization of the two prominent left-hemisphere peaks emerging from the lateralized cross-sectional SPM analysis (Fig. 2b) using the Anatomy toolbox. A spherical ROI (radius = 8 mm; 257 voxels) was centered at each peak, and mirrored in the right hemisphere. "Overlap" refers to the percentage of the sphere which overlapped with voxels in the associated maximum probability map (MPM). "Superior temporal" refers to those voxels in the AAL map for STG after masking out all voxels within the Anatomy toolbox MPM.

ROI	Region	Subdivision	Overlap with ROI (%)	
			Left	Right
Peak 1	Supramarginal	PF	15.3	3.4
		PFcm	46.1	19.9
		PFm	2.1	42.7
		–	26.5	9.2
Peak 2	Superior temporal	PGa	0.9	12.7
	Angular	TE 1.0	44.7	27.4
		TE 1.1	15.3	23.3
	Heschl's	–	28.9	44.5
		OP 1	4.8	4.7
	Superior temporal	–	–	–
	Operculum	–	–	–

**Table 3**

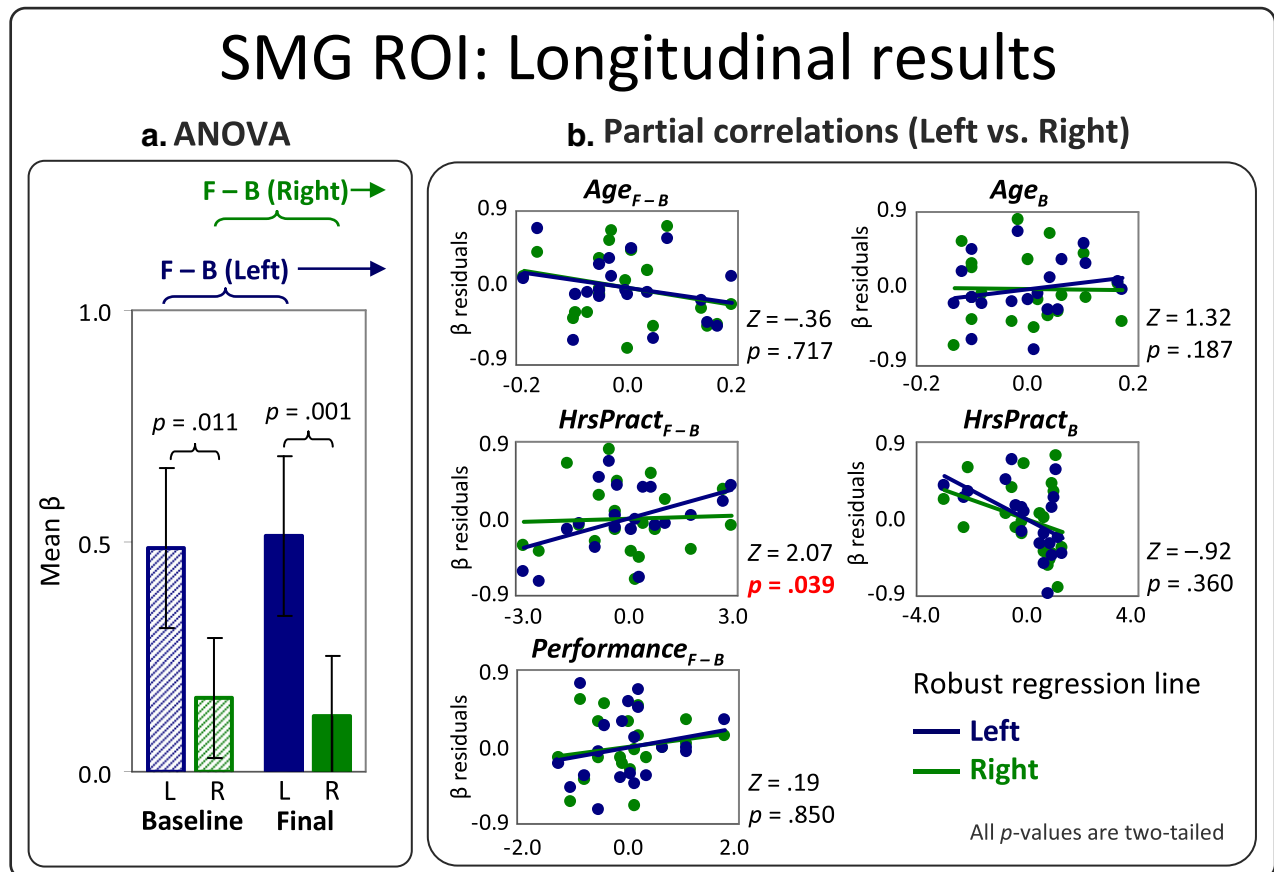
Summary of all partial correlations from the cross-sectional and longitudinal regressions. The key statistic of interest is  $Z$ , which indicates whether the correlation strength with a given predictor variable differed between the left and right hemispheres. For each regressor,  $L$  = left  $\beta$ -value residuals,  $R$  = right  $\beta$ -value residuals, and  $X$  = predictor variable residuals. Two-tailed  $p$ -values were calculated using  $(N - K - 1)$  degrees of freedom for correlations, and  $(N - K - 2)$  degrees of freedom for the  $Z$ -test.

Analysis	ROI	Predictor	Component correlations			Z-test
			$r(X,L)$	$r(X,R)$	$r(L,R)$	
Cross-sectional	SMG	$Age_S$	$r = -.090$	$r = -.063$	$r = .374$	$Z = -.22$
			$p = .422$	$p = .573$	$p < .001$	$p = .830$
		$HrsPract_S$	$r = .299$	$r = -.090$	$r = .323$	$Z = 2.99$
			$p = .006$	$p = .421$	$p = .003$	<b><math>p = .003</math></b>
		$Performance_S$	$r = -.069$	$r = .071$	$r = .360$	$Z = -1.10$
			$p = .540$	$p = .529$	$p < .001$	$p = .271$
	HG	$Age_S$	$r = -.025$	$r = .156$	$r = .473$	$Z = -1.57$
			$p = .824$	$p = .162$	$p < .001$	$p = .116$
		$HrsPract_S$	$r = -.120$	$r = -.169$	$r = .499$	$Z = .44$
Longitudinal	SMG		$p = .283$	$p = .129$	$p < .001$	$p = .659$
		$Performance_S$	$r = .010$	$r = -.032$	$r = .487$	$Z = .37$
			$p = .929$	$p = .775$	$p < .001$	$p = .713$
		$Age_{F-B}$	$r = -.296$	$r = -.227$	$r = .744$	$Z = -.36$
			$p = .265$	$p = .398$	$p = .001$	$p = .717$
		$HrsPract_{F-B}$	$r = .459$	$r = .003$	$r = .663$	$Z = 2.07$
			$p = .073$	$p = .991$	$p = .005$	<b><math>p = .039</math></b>
		$Performance_{F-B}$	$r = .170$	$r = .132$	$r = .730$	$Z = .19$
			$p = .529$	$p = .626$	$p = .001$	$p = .850$
		$Age_B$	$r = .238$	$r = -.047$	$r = .694$	$Z = 1.32$
			$p = .374$	$p = .863$	$p = .003$	$p = .187$
		$HrsPract_B$	$r = -.461$	$r = -.297$	$r = .751$	$Z = -.92$
			$p = .072$	$p = .264$	$p < .001$	$p = .360$

process musical versus linguistic stimuli (e.g., Specht et al., 2009; Tervaniemi and Hugdahl, 2003) or fine-grained differences in melodic versus rhythmic complexity (e.g., Schönwiesner et al., 2005; Zatorre and Belin, 2001); or when subjects perform speaking versus singing tasks (e.g., Suarez et al., 2010).

$pSTG$  vs. SMG

Before interpreting the observed correlation between musical training and SMG activation, it is necessary to briefly touch on the location of the current effect with respect to one from our previous



**Fig. 4.** Further exploration of SMG asymmetries in the longitudinal data set. (a) Mean  $\beta$ -values analyzed in a 2 (Time)  $\times$  2 (Hemisphere) ANOVA. (b) Robust regression results, visualizing the partial correlation of  $HrsPract_{F-B}$  (abscissa) with  $\beta$ -value residuals (ordinate) for each factor (separately in the left and right hemispheres), with the associated  $Z$ -test.

**Table 4**

Testing differences in partial correlation strength between  $HrsPract_S$  (“X”) and mean  $\beta$ -values during MD and RD in the cross-sectional data set; and the difference in left-hemisphere correlation strength between MD and RD.

Comparison (“Y <sub>1</sub> ” vs. “Y <sub>2</sub> ”)	Component correlations			Z-test
	$r(X, Y_1)$	$r(X, Y_2)$	$r(Y_1, Y_2)$	
MD <sub>L</sub> vs. MD <sub>R</sub>	$r = .328$ $p = .003$	$r = -.016$ $p = .886$	$r = .482$ $p < .001$	$Z = 3.05$ <b><math>p = .002</math></b>
RD <sub>L</sub> vs. RD <sub>R</sub>	$r = .281$ $p = .011$	$r = -.097$ $p = .386$	$r = .228$ $p = .039$	$Z = 2.72$ <b><math>p = .006</math></b>
MD <sub>L</sub> vs. RD <sub>L</sub>	$r = .328$ $p = .003$	$r = .281$ $p = .011$	$r = .587$ $p < .001$	$Z = .488$ $p = .625$

Significant  $p$ -values are highlighted in bold.

study (Ellis et al., 2012, Fig. 5). That is, using the same cross-sectional subjects, stimuli, and (non-lateralized) first-level contrasts, voxelwise regression analysis revealed a correlation with cumulative hours of training localized to left posterior superior temporal gyrus (pSTG)/planum temporale (PT). The peak of that cluster ( $[-60, -48, 12]$ ) is noticeably different from the peak labeled as SMG in the present analysis ( $[-52, -42, 22]$ ; a Euclidian distance of 14 mm). We suggest that these findings are not incompatible, as they represent the “ends” of rather different analytic means. The pSTG ROI in Ellis et al. (2012) was identified via a direct voxelwise regression on unlateralized contrasts that was independent of any overall task-related activity. In fact, a one-sample  $t$ -test on mean  $\beta$ -values extracted from the 248-voxel pSTG cluster was just below conventional significance ( $t_{83} = 1.96, p = .053$ ), and far below the threshold of detection in the whole-brain analysis ( $p_{\text{voxel}} < .001$ ). The present SMG ROI, on the other hand, was identified via a three-element conditional, visualized in Fig. 1d: (1) significant overall task-related activity in one hemisphere (here, left) that was (2) significantly lateralized (left > right) relative to homotopic voxels, and which (3) possessed a correlation with one of the regressors of interest that was also significantly lateralized (left > right).

Importantly, the lateralized statistical mapping technique (Conditional 2) does not account for interhemispheric differences in macroanatomy (e.g., Toga and Thompson, 2003); rather, it blurs such differences by using a symmetrical template during normalization. This has particular relevance for the present analysis, as our previous pSTG/PT cluster falls within a “hotspot” of significant positional displacement of cortical surface vertices between the left and right hemispheres (Lyttelton et al., 2009), a product of the anterior translocation of the posterior ascending ramus of the Sylvian fissure. This finding is not an issue for analyses based upon structural estimates of PT anatomy made separately within each hemisphere (e.g., Keenan et al., 2001; Schlaug et al., 1995b), or which functionally localize task-related activity in temporoparietal cortex separately for each subject (e.g., Hickok et al., 2003; Pa and Hickok, 2008). Using a lateralized SPM analysis to directly compare activity in homotopic voxels in the most posterior aspect of STG/PT, however, will likely present challenges. By contrast, the degree of overlap between the present functionally defined ROI and voxels belonging to cytoarchitecturally defined SMG (Caspers et al., 2008) was similar for the original ROI and its mirror image (~60%). SMG (Brodmann Area 40) was thus deemed a suitable label for the location of this lateralized effect.

Having localized the present effect to SMG, the remainder of the Discussion will center on interpreting the significance of our primary finding: a leftward asymmetry in overall task-related fMRI activity in SMG that was further characterized by a leftward asymmetry in correlation strength with inter- and intra-subject differences in subjects' musical training. Four themes will be presented: (1) left SMG and auditory working memory; (2) left SMG and (perceived) rhythmic complexity; (3) left SMG and temporally oriented attention; (4) left SMG and performance effects; and, more broadly, (5) age of commencement vs. duration of practice effects.

### Left SMG and auditory working memory

Left SMG has been implicated in the phonological short-term storage component of working memory (e.g., Baddeley, 2003; Paulesu et al., 1993; Romero et al., 2006; Vallar and Papagno, 2002), and the adjacent “Sylvian temporoparietal cortex” in the articulatory (vocal tract) rehearsal component of working memory (e.g., Buchsbaum et al., 2005; Hickok et al., 2003; Pa and Hickok, 2008). This distinction (phonological store versus articulatory loop) is beyond the scope of the present paradigm and analysis (but see Buchsbaum et al., 2011, for a recent review and synthesis). Instead, we highlight the importance of left SMG as a “general nodal point for short-term auditory working memory” (Vines et al., 2006, p. 1049) during operations on both verbal (phoneme sequences) and tonal (pitch sequences) stimuli (e.g., Gaab et al., 2003; Gaab et al., 2006; Koelsch et al., 2009; Schulze et al., 2011; Vines et al., 2006). That the pattern of results observed in SMG (i.e., a leftward asymmetry in task-related activity plus a leftward asymmetry in correlation strength with practice time) was not statistically different for melodic versus rhythmic discrimination is consistent with this broader view. It is also consistent with other reports of left SMG activity during the perception (e.g., Grahm et al., 2011; Janata et al., 2002; Limb et al., 2006), imagery (e.g., Zatorre et al., 1996), or overt motor production (Bengtsson and Ullén, 2006; Jungblut et al., 2012; Vuust et al., 2006) of both melodies and rhythms (cf. Ellis et al., 2012, Tables 1a and 1b). Our finding is also in nice agreement with a recent report by Herdener et al. (in press) in which professional jazz drummers were observed to engage left SMG more than non-musician matched controls when perceiving deviations from temporal regularity within complex rhythms.

### Left SMG and (perceived) rhythmic complexity

An alternative explanation for the observed leftward asymmetry in task-related activation asymmetry may be that it was driven not by working memory per se, but by perceiving the rhythmic nature of our stimuli. That is, our “melodic” stimuli could be described as “sequences of varying pitches with isochronous timing,” and our “rhythmic” stimuli as “sequences of a repeated pitch with a temporally complex timing.” Much has been written about relative asymmetries in cortical activation when healthy listeners perceive rhythmic (left > right) versus melodic (right > left) features of musical stimuli (e.g., Schönwiesner et al., 2005; Zatorre and Belin, 2001), bolstered by a large (albeit complex) corpus of lesion data (for reviews, see e.g., Peretz and Zatorre, 2005; Stewart et al., 2006).

If SMG were differentially responding to *perceived* rhythmic complexity (we will refrain from a hypothesis about *produced* rhythmic complexity, since our paradigm cannot address this), it could be predicted that the overall leftward asymmetry and/or degree of partial correlation strength with musical training should be stronger during the RD than the MD condition. A specific test of this hypothesis (cf. Table 4) did not achieve significance in the present data. It may well be the case, however, that our MD and RD stimuli were not different enough to elicit significant differences in SMG activation; perhaps a finer-grained manipulation of temporal complexity (e.g., Bengtsson et al., 2009 [4 levels]; Schönwiesner et al., 2005 [5 levels]) may have revealed significantly different partial correlations as a function of musical training.

### Left SMG and temporally oriented attention

Within a different literature, left SMG (as part of inferior parietal cortex) has been implicated (see Coull, 2004 for a review) in tasks requiring motor attention (i.e., cueing *which* motor action should occur; e.g., Rushworth et al., 2001) and the temporal orientation of attention (i.e., cueing *when* in time a motor action should occur; e.g., Coull and Nobre, 1998). A positive association between years of musical



training and increased recruitment of mechanisms supporting temporally oriented attention is both plausible and interesting. Importantly, however, the “limitations” of our own design (i.e., time and complexity constraints motivated by our youngest participants) precluded isolating passive listening and/or attentional processes from working memory and/or rehearsal processes (as in, e.g., Jerde et al., 2011; Schulze et al., 2011).

#### *Left SMG and performance effects*

The present results found practice-related, but not performance-related, effects on SMG activity. Several previous studies have reported the latter. Gaab et al. (2003) reported a positive correlation between subjects' performance and left SMG activity. This paradigm was expanded into a two time-point fMRI study by Gaab et al. (2006), in which subjects performed five days of pitch-memory training between the two scans. Contrasting subjects with larger (vs. smaller) performance improvements at the second (vs. the first) scan revealed focal activity in left anterior SMG. Additionally, Vines et al. (2006) found that transcranial direct current stimulation over left supramarginal gyrus (i.e., at the TP3 10–20 electrode site) resulted in poorer post-stimulation performance (change in percent correct) compared to sham stimulation or right SMG stimulation. None of these three studies, however, accounted for inter-subject differences in musical training (which ranged from elementary through collegiate level, although no subjects were active musicians at the time of the experiment). Furthermore, Gaab and Schlaug (2003) found that, when pitch-memory performance was matched between musicians and non-musicians (i.e., by selecting a sample of high-performing non-musicians, resulting in roughly 77% accuracy for both groups), a *Musicians > Non-musicians* contrast (i.e., a categorical effect) revealed primarily right SMG activity rather than left SMG activity.

Taken together, these studies suggest that performance-related effects in left SMG may be analysis dependent: when musical training and performance is stratified in the sample, and training differences are not taken into consideration (Gaab et al., 2003, 2006; Vines et al., 2006), performance-related effects may appear. By contrast, when both training and performance are modeled simultaneously (as in the present study; cf. Table 3), training-related effects emerge more strongly than performance-related effects in left SMG. Using conjunction analysis, Schulze et al. (2011) identified left SMG (i.e., ventral inferior parietal lobule) as one of several regions active when both musicians and non-musicians engage verbal and tonal working memory. The present results add to these previous findings, suggesting that left SMG shows a nuanced response that is modulated by inter- and intra-individual differences in musical training.

#### *Age of commencement vs. duration of practice*

As noted in the **Materials and methods**, the present design modeled the cumulative hours of practice but not the age of commencement of training. Although cumulative hours can only be treated as an approximation for the cross-sectional sample, it was deemed a more accurate index than a single “practice intensity” value (i.e., hours per unit of time, as in Hutchinson et al., 2003), which likely varied over the course of many years of training.

Several previous functional, structural, and electrophysiological studies have reported significant correlations (across the brain) with age of commencement (without evaluating cumulative practice; e.g., Amunts et al., 1997; Keenan et al., 2001; Pantev et al., 1998; Schlaug et al., 1995a); others have reported significant correlations with cumulative practice (without evaluating age of commencement; e.g., Bengtsson et al., 2005; Kleber et al., 2010; Musacchia et al., 2007). Studies which have simultaneously analyzed both variables have yielded heterogeneous results: an effect of practice intensity but not years of practice or age of commencement (Hutchinson et al., 2003); an effect of age of commencement but not cumulative practice

(Ohnishi et al., 2001); an effect of cumulative practice but not age of commencement (Elmer et al., 2012; Sluming et al., 2002); and effects for both age of commencement and cumulative practice, but stronger for the latter (Foster and Zatorre, 2010).

Certainly, differences in task (e.g., perceptual vs. motor) and imaging modality (structural vs. functional) will play a strong role in these differences. Just as (statistically) important, however, is an accurate sampling across the plausible range of values for each dimension (cf. Bland and Altman, 2011): including subjects with a range of commencement ages, total years of practice, and intensity of practice. (Depending on the behavioral task, these subject differences should translate into a range of performance levels, which will further play into the resultant pattern of effect significance.) Additionally, it could be hypothesized that lengthening the longitudinal timeframe would translate into greater differences in SMG functional lateralization, mirroring the cross-sectional pattern of results (cf. Fig. 3d). Although it was not possible to evaluate age of commencement in the present analysis without compromising statistical validity, it was possible to statistically isolate musical training over the course of the study (between Baseline and Final) from musical training prior to Baseline; as well as from the “general” effect of maturation (elapsed years between Baseline and Final). While this does not directly capture the age at which training commenced, it does move the analysis one step closer towards the simultaneous assessment of multiple aspects of musical training that may impact perceptual processing and its neural correlates.

#### **Conclusion**

In combining the sensitivity of a large cross-sectional design with the specificity of a smaller longitudinal design, the present analyses help clarify how musical training shapes patterns of cortical activity when listeners discriminate musical stimuli. Both inter- and intra-subject differences in training were associated with a leftward asymmetry in activity in (ventral) supramarginal gyrus during melodic and rhythmic discrimination. This suggests a direct link between the amount of time spent practicing an instrument and the degree of engagement of working memory mechanisms, and reinforces the importance of this region in auditory working memory more broadly.

#### **Software note**

Several of the computational steps involved in creating lateralized contrasts were performed using custom-built, SPM-compatible MATLAB scripts, available for download at <http://tools.robjellis.net>.

#### **Conflict of interest statement**

No conflicts of interest are reported.

#### **Acknowledgment**

Many thanks to our co-investigator Katie Overy; to the many research assistants (C. Alexander, L. Blake, K. Brumm, K. Cronin, K. Farrell, L. Forbes, M. Forgeard, U. Iyengar, M. Rosam, and L. Zhu) for test preparation, behavioral testing, and imaging data collection; and to two anonymous reviewers for insightful comments and suggestions. This work was supported by grants from the National Science Foundation (BCS-0132508 and BCS-0518837) and supplemental funds from the Dana Foundation and the Grammy Foundation. Additional support was provided by grants from the National Institutes of Health (DC009823 and DC008796) to G.S., and from the Harvard NeuroDiscovery Center to R.J.E.

## References

- Abbott, D.F., Waites, A.B., Lillywhite, L.M., Jackson, G.D., 2010. fMRI assessment of language lateralization: an objective approach. *NeuroImage* 50 (4), 1446–1455.
- Amunts, K., Schlaug, G., Jäncke, L., Steinmetz, H., Schleicher, A., Dabringhaus, A., Zilles, K., 1997. Motor cortex and hand motor skills: structural compliance in the human brain. *Hum. Brain Mapp.* 5 (3), 206–215.
- Annett, M., 1970. A classification of hand preference by association analysis. *Br. J. Psychol.* 61 (3), 303–321.
- Baddeley, A., 2003. Working memory: looking back and looking forward. *Nat. Rev. Neurosci.* 4 (10), 829–839.
- Bangert, M., Altenmüller, E.O., 2003. Mapping perception to action in piano practice: a longitudinal DC-EEG study. *BMC Neurosci.* 4, 26.
- Belin, P., Zatorre, R.J., Hoge, R., Evans, A.C., Pike, B., 1999. Event-related fMRI of the auditory cortex. *NeuroImage* 10 (4), 417–429.
- Bengtsson, S.L., Ullén, F., 2006. Dissociation between melodic and rhythmic processing during piano performance from musical scores. *NeuroImage* 30 (1), 272–284.
- Bengtsson, S., Nagy, Z., Skare, S., Forsman, L., Forssberg, H., Ullén, F., 2005. Extensive piano practicing has regionally specific effects on white matter development. *Nat. Neurosci.* 8 (9), 1148–1150.
- Bengtsson, S.L., Ullén, F., Ehrsson, H.H., Hashimoto, T., Kito, T., Naito, E., et al., 2009. Listening to rhythms activates motor and premotor cortices. *Cortex* 45, 62–71.
- Besson, M., Chobert, J., Marie, C., 2011. Transfer of training between music and speech: common processing, attention, and memory. *Front. Psychol.* 2, 94.
- Bever, T.G., Chiarello, R.J., 1974. Cerebral dominance in musicians and nonmusicians. *Science* 185 (4150), 537–539.
- Bland, J.M., Altman, D.G., 2011. Correlation in restricted ranges of data. *BMJ* 342, d556.
- Büchel, C., Raedler, T., Sommer, M., Sach, M., Weiller, C., Koch, M.A., 2004. White matter asymmetry in the human brain: a diffusion tensor MRI study. *Cereb. Cortex* 14 (9), 945–951.
- Buchsbaum, B.R., Olsen, R.K., Koch, P., Berman, K.F., 2005. Human dorsal and ventral auditory streams subserve rehearsal-based and echoic processes during verbal working memory. *Neuron* 48 (4), 687–697.
- Buchsbaum, B.R., Baldo, J., Okada, K., Berman, K.F., Dronkers, N., D'Esposito, M., Hickok, G., 2011. Conduction aphasia, sensory-motor integration, and phonological short-term memory—an aggregate analysis of lesion and fMRI data. *Brain Lang.* 119 (3), 119–128.
- Caspers, S., Eickhoff, S.B., Geyer, S., Scheperjans, F., Mohlberg, H., Zilles, K., Amunts, K., 2008. The human inferior parietal lobule in stereotaxic space. *Brain Struct. Funct.* 212 (6), 481–495.
- Coull, J.T., 2004. fMRI studies of temporal attention: allocating attention within, or towards, time. *Cogn. Brain Res.* 21 (2), 216–226.
- Coull, J.T., Nobre, A.C., 1998. Where and when to pay attention: the neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. *J. Neurosci.* 18 (18), 7426–7435.
- Eickhoff, S.B., Stephan, K.E., Mohlberg, H., Grefkes, C., Fink, G.R., Amunts, K., Zilles, K., 2005. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage* 25 (4), 1325–1335.
- Ellis, R.J., Norton, A.C., Overy, K., Winner, E., Alsop, D.C., Schlaug, G., 2012. Differentiating maturational and training influences on fMRI activation during music processing. *NeuroImage* 60 (3), 1902–1912.
- Elmer, S., Meyer, M., Jäncke, L., 2012. Neurofunctional and behavioral correlates of phonetic and temporal categorization in musically trained and untrained subjects. *Cereb. Cortex* 22 (3), 650–658.
- Forgeard, M., Winner, E., Norton, A., Schlaug, G., 2008. Practicing a musical instrument in childhood is associated with enhanced verbal ability and nonverbal reasoning. *PLoS One* 3 (10), e3566.
- Foster, N.E.V., Zatorre, R.J., 2010. Cortical structure predicts success in performing musical transformation judgments. *NeuroImage* 53 (1), 26–36.
- Fujioka, T., Ross, B., Kakigi, R., Pantev, C., Trainor, L., 2006. One year of musical training affects development of auditory cortical-evoked fields in young children. *Brain* 129, 2593–2608.
- Gaab, N., Schlaug, G., 2003. Musicians differ from nonmusicians in brain activation despite performance matching. *Ann. N.Y. Acad. Sci.* 999, 385–388.
- Gaab, N., Gaser, C., Zaehle, T., Jäncke, L., Schlaug, G., 2003. Functional anatomy of pitch memory—an fMRI study with sparse temporal sampling. *NeuroImage* 19 (4), 1417–1426.
- Gaab, N., Gaser, C., Schlaug, G., 2006. Improvement-related functional plasticity following pitch memory training. *NeuroImage* 31 (1), 255–263.
- Grahn, J.A., Henry, M.J., McAuley, J.D., 2011. fMRI investigation of cross-modal interactions in beat perception: audition primes vision, but not vice versa. *NeuroImage* 54 (2), 1231–1243.
- Griffiths, T.D., Warren, J.D., 2002. The planum temporale as a computational hub. *Trends Neurosci.* 25 (7), 348–353.
- Herdener, M., Esposito, F., di Salle, F., Boller, C., Hilti, C.C., Habermeyer, B., Scheffler, K., et al., 2010. Musical training induces functional plasticity in human hippocampus. *J. Neurosci.* 30 (4), 1377.
- Herdener, M., Humbel, T., Esposito, F., Habermeyer, B., Cattapan-Ludewig, K., Seifritz, E., in press. Jazz Drummers Recruit Language-Specific Areas for the Processing of Rhythmic Structure. <http://dx.doi.org/10.1093/cercor/bhs367>.
- Herholz, S.C., Boh, B., Pantev, C., 2011. Musical training modulates encoding of higher-order regularities in the auditory cortex. *Eur. J. Neurosci.* 34 (3), 524–529.
- Hickok, G., Buchsbaum, B., Humphries, C., Muftuler, T., 2003. Auditory-motor interaction revealed by fMRI: speech, music, and working memory in area Spt. *J. Cogn. Neurosci.* 15 (5), 673–682.
- Ho, Y.C., Cheung, M.C., Chan, A.S., 2003. Music training improves verbal but not visual memory: cross-sectional and longitudinal explorations in children. *Neuropsychology* 17 (3), 439–450.
- Hutchinson, S., Lee, L.H.-L., Gaab, N., Schlaug, G., 2003. Cerebellar volume of musicians. *Cereb. Cortex* 13 (9), 943–949.
- Hyde, K.L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A.C., Schlaug, G., 2009. Musical training shapes structural brain development. *J. Neurosci.* 29 (10), 3019–3025.
- Jabusch, H.C., Alpers, H., Kopiez, R., Vauth, H., Altenmüller, E., 2009. The influence of practice on the development of motor skills in pianists: a longitudinal study in a selected motor task. *Hum. Mov. Sci.* 28 (1), 74–84.
- Janata, P., Tillmann, B., Bharucha, J.J., 2002. Listening to polyphonic music recruits domain-general attention and working memory circuits. *Cogn. Affect. Behav. Neurosci.* 2 (2), 121–140.
- Jäncke, L., 2009. The plastic human brain. *Restor. Neurol. Neurosci.* 27 (5), 521–538.
- Jansen, A., Menke, R., Sommer, J., Förster, A.F., Bruchmann, S., Hempleman, J., Weber, B., et al., 2006. The assessment of hemispheric lateralization in functional MRI—robustness and reproducibility. *NeuroImage* 33 (1), 204–217.
- Jerde, T.A., Childs, S.K., Handy, S.T., Nagode, J.C., Pardo, J.V., 2011. Dissociable systems of working memory for rhythm and melody. *NeuroImage* 57 (4), 1572–1579.
- Jungblut, M., Huber, W., Pustelniak, M., Schnitker, R., 2012. The impact of rhythm complexity on brain activation during simple singing: an event-related fMRI study. *Restor. Neurol. Neurosci.* 30 (1), 39–53.
- Keenan, J.P., Thangaraj, V., Halpern, A.R., Schlaug, G., 2001. Absolute pitch and planum temporale. *NeuroImage* 14 (6), 1402–1408.
- Keppel, G., Wickens, T.D., 2004. Design and Analysis: A Researcher's Handbook, 4th ed. Prentice Hall, Upper Saddle River, NJ.
- Kleber, B., Veit, R., Birbaumer, N., Gruzelier, J., Lotze, M., 2010. The brain of opera singers: experience-dependent changes in functional activation. *Cereb. Cortex* 20 (5), 1144–1152.
- Koelsch, S., Schulze, K., Sammler, D., Fritz, T., Müller, K., Gruber, O., 2009. Functional architecture of verbal and tonal working memory: an fMRI study. *Hum. Brain Mapp.* 30 (3), 859–873.
- Kutner, M., Nachtsheim, C., Neter, J., 2004. Applied Linear Regression Models, 4th ed. McGraw-Hill.
- Liégeois, F., Connelly, A., Salmond, C.H., Gadian, D.G., Vargha-Khadem, F., Baldeweg, T., 2002. A direct test for lateralization of language activation using fMRI: comparison with invasive assessments in children with epilepsy. *NeuroImage* 17 (4), 1861–1867.
- Limb, C.J., Kemény, S., Ortigoza, E.B., Rouhani, S., Braun, A.R., 2006. Left hemispheric lateralization of brain activity during passive rhythm perception in musicians. *Anat. Rec. A Discov. Mol. Cell. Evol. Biol.* 288 (4), 382–389.
- Luders, E., Gaser, C., Jancke, L., Schlaug, G., 2004. A voxel-based approach to gray matter asymmetries. *NeuroImage* 22 (2), 656–664.
- Lyttelton, O.C., Karama, S., Ad-Dabagh, Y., Zatorre, R.J., Carbonell, F., Worsley, K., Evans, A.C., 2009. Positional and surface area asymmetry of the human cerebral cortex. *NeuroImage* 46 (4), 895–903.
- MacCallum, R.C., Zhang, S., Preacher, K.J., Rucker, D.D., 2002. On the practice of dichotomization of quantitative variables. *Psychol. Methods* 7 (1), 19–40.
- Macmillan, N.A., Creelman, C.D., 2005. Detection Theory: A User's Guide. Lawrence Erlbaum.
- Meng, X.L., Rosenthal, R., Rubin, D.B., 1992. Comparing correlated correlation coefficients. *Psychol. Bull.* 111 (1), 172–175.
- Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S.L., Besson, M., 2009. Musical training influences linguistic abilities in 8-year-old children: more evidence for brain plasticity. *Cereb. Cortex* 19 (3), 712–723.
- Musacchia, G., Sams, M., Skoe, E., Kraus, N., 2007. Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proc. Natl. Acad. Sci. U. S. A.* 104 (40), 15894–15898.
- Nichols, T.E., 2008. CorrClustTh. Retrieved from <http://www2.warwick.ac.uk/fac/sci/statistics/staff/academic-research/nichols/scripts/spm/johnsgems5/CorrClustTh.m>.
- Norton, A., Winner, E., Cronin, K., Overy, K., Lee, D.J., Schlaug, G., 2005. Are there pre-existing neural, cognitive, or motoric markers for musical ability? *Brain Cogn.* 59 (2), 124–134.
- Ohnishi, T., Matsuda, H., Asada, T., Aruga, M., Hirakata, M., Nishikawa, M., Katoh, A., et al., 2001. Functional anatomy of musical perception in musicians. *Cereb. Cortex* 11 (8), 754–760.
- Overy, K., Norton, A.C., Cronin, K.T., Gaab, N., Alsop, D.C., Winner, E., Schlaug, G., 2004. Imaging melody and rhythm processing in young children. *Neuroreport* 15 (11), 1723–1726.
- Pa, J., Hickok, G., 2008. A parietal-temporal sensory-motor integration area for the human vocal tract: evidence from an fMRI study of skilled musicians. *Neuropsychologia* 46 (1), 362–368.
- Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L.E., Hoke, M., 1998. Increased auditory cortical representation in musicians. *Nature* 392 (6678), 811–814.
- Paulesu, E., Frith, C.D., Frackowiak, R.S., 1993. The neural correlates of the verbal component of working memory. *Nature* 362 (6418), 342–345.
- Penhune, V.B., 2011. Sensitive periods in human development: evidence from musical training. *Cortex* 47 (9), 1126–1137.
- Peretz, I., Zatorre, R.J., 2005. Brain organization for music processing. *Annu. Rev. Psychol.* 56, 89–114.
- Poldrack, R.A., 2007. Region of interest analysis for fMRI. *Soc. Cogn. Affect. Neurosci.* 2, 67–70.
- Poldrack, R.A., Mumford, J.A., Nichols, T.E., 2011. Handbook of Functional MRI Data Analysis. Cambridge University Press, New York.
- Romero, L., Walsh, V., Papagno, C., 2006. The neural correlates of phonological short-term memory: a repetitive transcranial magnetic stimulation study. *J. Cogn. Neurosci.* 18 (7), 1147–1155.
- Rushworth, M.F., Ellison, A., Walsh, V., 2001. Complementary localization and lateralization of orienting and motor attention. *Nat. Neurosci.* 4 (6), 656–661.

- Schellenberg, E.G., 2004. Music lessons enhance IQ. *Psychol. Sci.* 15 (8), 511.
- Schlaug, G., Jäncke, L., Huang, Y., Steinmetz, H., 1995a. In vivo evidence of structural brain asymmetry in musicians. *Science* 267 (5198), 699–701.
- Schlaug, G., Jäncke, L., Huang, Y., Staiger, J.F., Steinmetz, H., 1995b. Increased corpus callosum size in musicians. *Neuropsychologia* 33 (8), 1047–1055.
- Schlaug, G., Norton, A., Overy, K., Winner, E., 2005. Effects of music training on the child's brain and cognitive development. *Ann. N.Y. Acad. Sci.* 1060, 219–230.
- Schlaug, G., Forgeard, M., Zhu, L., Norton, A., Norton, A., Winner, E., 2009. Training-induced neuroplasticity in young children. *Ann. N.Y. Acad. Sci.* 1169, 205–208.
- Schönwiesner, M., Rübsem, R., von Cramon, D.Y., 2005. Hemispheric asymmetry for spectral and temporal processing in the human antero-lateral auditory belt cortex. *Eur. J. Neurosci.* 22 (6), 1521–1528.
- Schulze, K., Zysset, S., Mueller, K., Friederici, A.D., Koelsch, S., 2011. Neuroarchitecture of verbal and tonal working memory in nonmusicians and musicians. *Hum. Brain Mapp.* 32 (5), 771–783.
- Sluming, V., Barrick, T., Howard, M., Cezayirli, E., Mayes, A., Roberts, N., 2002. Voxel-based morphometry reveals increased gray matter density in Broca's area in male symphony orchestra musicians. *NeuroImage* 17 (3), 1613–1622.
- Specht, K., Osnes, B., Hugdahl, K., 2009. Detection of differential speech-specific processes in the temporal lobe using fMRI and a dynamic “sound morphing” technique. *Hum. Brain Mapp.* 30 (10), 3436–3444.
- Stevens, M.C., Calhoun, V.D., Kiehl, K.A., 2005. Hemispheric differences in hemodynamics elicited by auditory oddball stimuli. *NeuroImage* 26 (3), 782–792.
- Stewart, L., Von Kriegstein, K., Warren, J.D., Griffiths, T.D., 2006. Music and the brain: disorders of musical listening. *Brain* 129 (10), 2533–2553.
- Suarez, R.O., Golby, A., Whalen, S., Sato, S., Theodore, W.H., Kufra, C.V., Devinsky, O., et al., 2010. Contributions to singing ability by the posterior portion of the superior temporal gyrus of the non-language-dominant hemisphere. *Cortex* 46 (3), 343–353.
- Swanson, N., Eichele, T., Pearlson, G., Kiehl, K., Yu, Q., Calhoun, V.D., 2011. Lateral differences in the default mode network in healthy controls and patients with schizophrenia. *Hum. Brain Mapp.* 32 (4), 654–664.
- Tervaniemi, M., Hugdahl, K., 2003. Lateralization of auditory-cortex functions. *Brain Res. Rev.* 43 (3), 231–246.
- Tervaniemi, M., Sannemann, C., Noyranen, M., Salonen, J., Pihko, E., 2011. Importance of the left auditory areas in chord discrimination in music experts as demonstrated by MEG. *Eur. J. Neuro.* 34 (3), 517–523.
- Toga, A.W., Thompson, P.M., 2003. Mapping brain asymmetry. *Nat. Rev. Neurosci.* 4 (1), 37–48.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., et al., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage* 15 (1), 273–289.
- Vallar, G., Papagno, C., 2002. Neuropsychological impairments of short-term memory. In: Baddeley, A.D., Kopelman, M.D., Wilson, B.A. (Eds.), *Handbook of Memory Disorders*. Wiley, Chichester, pp. 249–270.
- Vines, B.W., Schneider, N.M., Schlaug, G., 2006. Testing for causality with transcranial direct current stimulation: pitch memory and the left supramarginal gyrus. *Neuroreport* 17 (10), 1047–1050.
- Vuust, P., Roepstorff, A., Wallentin, M., Mouridsen, K., Østergaard, L., 2006. It don't mean a thing. Keeping the rhythm during polyrhythmic tension, activates language areas (BA47). *NeuroImage* 31 (2), 832–841.
- Wager, T.D., Keller, M.C., Lacey, S.C., Jonides, J., 2005. Increased sensitivity in neuroimaging analyses using robust regression. *NeuroImage* 26 (1), 99–113.
- Wan, C.Y., Schlaug, G., 2010. Music making as a tool for promoting brain plasticity across the life span. *Neuroscientist* 16 (5), 566–577.
- Warren, J.E., Wise, R.J.S., Warren, J.D., 2005. Sounds do-able: auditory-motor transformations and the posterior temporal plane. *Trends Neurosci.* 28 (12), 636–643.
- Yarkoni, T., 2009. Big correlations in little studies: Inflated fMRI correlations reflect low statistical power—commentary on Vul et al. (2009). *Perspect. Psychol. Sci.* 4 (3), 294–298.
- Zatorre, R.J., Belin, P., 2001. Spectral and temporal processing in human auditory cortex. *Cereb. Cortex* 11 (10), 946–953.
- Zatorre, Robert J., Halpern, A.R., Perry, D.W., Meyer, E., Evans, A.C., 1996. Hearing in the mind's ear: a pet investigation of musical imagery and perception. *J. Cogn. Neurosci.* 8, 29–46.
- Zatorre, Robert J., Chen, J.L., Penhune, V.B., 2007. When the brain plays music: auditory-motor interactions in music perception and production. *Nat. Rev. Neurosci.* 8 (7), 547–558.