

Brain Activation during Anticipation of Sound Sequences

Amber M. Leaver,¹ Jennifer Van Lare,¹ Brandon Zielinski,¹ Andrea R. Halpern,² and Josef P. Rauschecker¹

¹Laboratory of Integrative Neuroscience and Cognition, Department of Physiology and Biophysics, Georgetown University Medical Center, Washington, DC 20057, and ²Psychology Department, Bucknell University, Lewisburg, Pennsylvania 17837

Music consists of sound sequences that require integration over time. As we become familiar with music, associations between notes, melodies, and entire symphonic movements become stronger and more complex. These associations can become so tight that, for example, hearing the end of one album track can elicit a robust image of the upcoming track while anticipating it in total silence. Here, we study this predictive “anticipatory imagery” at various stages throughout learning and investigate activity changes in corresponding neural structures using functional magnetic resonance imaging. Anticipatory imagery (in silence) for highly familiar naturalistic music was accompanied by pronounced activity in rostral prefrontal cortex (PFC) and premotor areas. Examining changes in the neural bases of anticipatory imagery during two stages of learning conditional associations between simple melodies, however, demonstrates the importance of fronto-striatal connections, consistent with a role of the basal ganglia in “training” frontal cortex (Pasupathy and Miller, 2005). Another striking change in neural resources during learning was a shift between caudal PFC earlier to rostral PFC later in learning. Our findings regarding musical anticipation and sound sequence learning are highly compatible with studies of motor sequence learning, suggesting common predictive mechanisms in both domains.

Key words: prefrontal cortex; basal ganglia; auditory; fMRI; learning and memory; motor learning

Introduction

When two melodies are frequently heard in the same order, as with consecutive movements of a symphony or tracks on a rock album, the beginning of the second melody is often anticipated vividly during the silence following the first. This reflexive, often irrepressible, retrieval of the second melody, or “anticipatory imagery,” reveals that music consists of cued associations, in this case between entire melodies. Storage of these complex associations can span seconds or minutes and therefore cannot be accomplished by single neurons on the basis of, for example, temporal combination sensitivity (Suga et al., 1978; Margoliash and Fortune, 1992; Rauschecker et al., 1995). Alternative neural mechanisms must enable the encoding and integration of sound sequences over longer intervals.

Like music processing, other forms of behavior require the integration and storage of temporal sequential hierarchies as well. This includes, for instance, speech perception, dancing, and athletic performance. Thus, some of the neural mechanisms underlying these behaviors may also be shared, and one may postulate the existence of generalized “sequence” areas that interact with auditory areas to facilitate learning of complex auditory sequences (Janata and Grafton, 2003; Schubotz and von Cramon, 2003; Zatorre et al., 2007).

How specific musical sequences are encoded and stored in the

brain is a neglected area of research. The few studies that compare neural responses to familiar and unfamiliar music (Platel et al., 2003; Plailly et al., 2007; Watanabe et al., 2008) only address recognition memory and do not tap other important memory processes such as recall and cued retrieval. Other studies that do address recall or retrieval confound the issue by requiring simultaneous execution of motor programs such as singing or playing an instrument (Sergent et al., 1992; Perry et al., 1999; Callan et al., 2006; Lahav et al., 2007). Imagery research, in contrast, can assess the neural mechanisms of music recall independent of concurrent sensorimotor events. Imagery for isolated sounds recruits nonprimary auditory cortex and the supplementary motor area, suggesting these areas are critical for the internal sensation of sound (Halpern et al., 2004). Imagery for melodies, including cued (mental) completion of familiar melodies, is associated with a wider network of frontal and parietal regions (Zatorre et al., 1996; Halpern and Zatorre, 1999). However, it is unknown how the complex, predictive cued associations that drive anticipatory imagery, as described above, are stored, or how these representations change with learning.

In our study, we exploit the anticipatory imagery phenomenon to assess how complex associations between musical sequences are learned and stored in the brain. Using functional magnetic resonance imaging (fMRI), we measured neural activity associated with anticipatory retrieval of a melody cued by its learned association with a previously unrelated melody, in two complementary learning situations. Experiment 1 assessed long-term incidental memorization of associated pieces of music. In experiment 2, we targeted anticipatory imagery during short-term, conscious memorization of simple melody pairs at two stages of this short-term training. Thus, we tracked how cued

Received Oct. 2, 2008; revised Dec. 30, 2008; accepted Jan. 3, 2009.

This work was supported by National Institutes of Health Grants R01-DC03489 and NS052494 (J.P.R.), F31-MH12598 (B.Z.), and F31-DC008921 (A.M.L.) and by the Cognitive Neuroscience Initiative of the National Science Foundation (Grants BCS-0350041 and BCS-0519127 to J.P.R. and a Research Opportunity Award to A.R.H.).

Correspondence should be addressed to Josef P. Rauschecker, Georgetown University Medical Center, 3970 Reservoir Road NW, Washington, DC 20057. E-mail: rauschej@georgetown.edu.

DOI:10.1523/JNEUROSCI.4921-08.2009

Copyright © 2009 Society for Neuroscience 0270-6474/09/292477-09\$15.00/0

associations between melodies are encoded throughout different stages of learning.

Materials and Methods

Participants

Twenty neurologically normal volunteers (10 males; mean age, 24.7 years) were recruited from the Georgetown University community to participate in these experiments. All were native English speakers, reported normal hearing, and were fully informed verbally and in writing of the procedures involved. Nine subjects (three females) participated in experiment 1; 11 subjects (seven females) participated in experiment 2, although one participant was omitted because of excessive head movement and poor behavioral performance. Participants in experiment 2 had at least 2 years of musical experience (mean, 6.5; SD, 4.17) in musical ensembles, lessons, or serious independent study.

Stimuli

Experiment 1 stimuli were digital recordings of the final 32 s of instrumental compact disk (CD) tracks that were either highly familiar or unfamiliar to a given subject. Experiment 2 stimuli consisted of 36 melodies constructed by one of the authors. Each was 2.5 s long, written in either C or F major, ended on either the tonic or the fifth scale degree, presented in a synthetic piano timbre, and judged as highly musical by two independent, moderately trained amateur musicians (rating of at least 7 on a 10-point scale, with 10 meaning very musical and 1 meaning not at all musical). In both experiments, all stimuli were gain-normalized and presented binaurally at a comfortable volume (~75–80 dB sound pressure level). During scans, sounds were delivered via a custom air-conduction sound system fitted with silicone-cushioned headphones specifically designed to isolate the subject from scanner noise (Magnetic Resonance Technologies).

Stimulus paradigm

Experiment 1

During experiment 1 scans, trials consisted of 32 s of music (the final seconds of familiar or unfamiliar CD tracks), followed by an 8 s period of silence. The order of music track presentation was randomized throughout each functional run. Each run consisted of nine trials, preceded by 16 s of silence during which two preexperimental baseline echo planar image (EPI) volumes were acquired (discarded in further analysis). Stimuli presented during familiar trials were the final 32 s of tracks from their favorite CD [familiar music (FM)], in which participants reported experiencing anticipatory imagery for each subsequent track in the silent gaps between music tracks [anticipatory silence (AS)]. Performance on next-track anticipation (singing or humming of the opening bars of the correct next track; 75% threshold for inclusion) for familiar music was confirmed by prescan behavioral testing. Stimuli presented during unfamiliar trials consisted of music that the subjects had never heard before [unfamiliar music (UM)]. Thus, during this condition, subjects would not anticipate the onset of the following track [nonanticipatory silence (NS)]. Three “familiar” functional runs were alternated with three “unfamiliar” functional runs. In total, 270 functional volumes were collected for each subject, yielding a total session time of approximately 1 h per subject. While in the magnetic resonance scanner, subjects were instructed to attend to the stimulus being presented and to imagine, but not vocalize, the subsequent melody where appropriate.

Experiment 2

Behavioral training. Before scanning in experiment 2, participants completed 30 min of self-directed training, in which they were asked to memorize seven pairs of melodies. Participants were told that the goal of training was to be able to retrieve and imagine the second melody accurately and vividly when they heard the corresponding first melody. Learning was assessed during a two-part testing phase. In the first phase, participants completed a discrimination task, in which melody pairs were presented and they were asked to detect a musically valid, one-note change in melody 2. The second phase was a recall task. Here, participants heard the first melody of a pair followed by silence, during which they imagined the second melody. The second melody was then pre-

sented, and participants rated their image on accuracy and vividness. In total, this training session lasted ~45 min. Participants then completed an additional ~10 min of self-directed “refresher” training before the scanning session and ~7–8 min more training between scan runs. Assigned trained and untrained melodies were counterbalanced across subjects.

Behavioral task. During scan trials in experiment 2, participants heard a familiar melody (i.e., the first melody of a trained pair), an unfamiliar melody, or silence. Participants were not cued as to the content of upcoming trials and therefore assessed the condition tested in each trial based on the music presented, or lack thereof. In the silence following familiar music, participants were asked to imagine the corresponding second melody of the trained pair (AS). In the silence following unfamiliar music, participants simply waited until the trial’s end (NS). At the end of each trial, participants rated the vividness of the image, if any, on a 5-point scale (1, no image; 5, very vivid image), regardless of trial type. In addition to assessing the vividness of the musical imagery experienced by the participants, these ratings also served to assess learning in runs 1 and 2. There were 70 trials with familiar melodies, 70 unfamiliar trials, and 34 silent trials per run, for two runs. Because of time restrictions, one participant heard 44 familiar, 44 unfamiliar, and 22 silent baseline trials per run, for two runs.

MRI protocol

Experiment 1

In experiment 1, images were obtained using a 1.5-tesla Siemens Magnetom Vision whole-body scanner. Functional EPIs were separated by several seconds of silence in a sparse sampling paradigm, allowing for a lesser degree of contamination of the EPIs and stimulus presentation by scanner noise (Hall et al., 1999). The timing of stimulus presentation relative to EPI acquisition was jittered to capture activation associated with either music (FM and UM) or silence after music (AS and NS). Functional runs consisted of 47 volumes each [repetition time (TR), 8 s; acquisition time (TA), ~2 s; echo time (TE), 40 ms; 25 axial slices; voxel size, $3.5 \times 3.5 \times 4.0 \text{ mm}^3$]. Coplanar high-resolution anatomical images were acquired using a T2-weighted sequence (TR, 5 s; TE, 99 ms; field of view, 240 mm; matrix, 192×256 ; voxel size, $0.94 \times 0.94 \times 4.4 \text{ mm}^3$).

Experiment 2

In experiment 2, images were acquired using a 3-tesla Siemens Trio scanner. Two sets of functional EPIs were acquired using a sparse sampling paradigm similar to experiment 1 (TR, 10 s; TA, 2.28 ms; TE, 30 ms; 50 axial slices; $3 \times 3 \times 3 \text{ mm}^3$ resolution). Between runs, a T1-weighted anatomical scan ($1 \times 1 \times 1 \text{ mm}^3$ resolution) was performed.

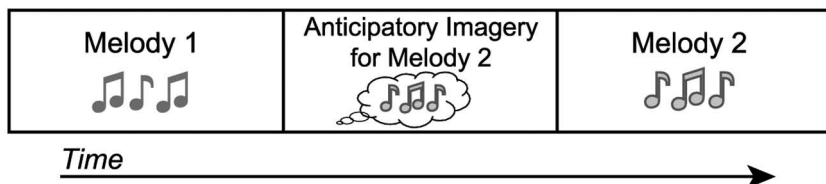
fMRI analysis

All analyses were completed using BrainVoyager QX (Brain Innovation). Functional images from each run were corrected for motion in six directions, smoothed using a three-dimensional 8 mm^3 Gaussian filter, corrected for linear trends, and high-pass filtered to remove low-frequency noise. Processed functional data were coregistered with corresponding high-resolution anatomical images and interpolated into Talairach space (Talairach and Tournoux, 1988).

Random-effects group analyses were performed across the entire brain and in regions of interest (ROIs) using the general linear model (GLM), which assessed the extent to which variation in blood oxygenation-dependent (BOLD)–fMRI signal can be explained by our experimental manipulations (i.e., predictors or regressors) (Friston et al., 1995). Random-effects models attempt to remove intersubject variability, thereby making the results obtained from a limited sample more generalizable to the entire population (Peterson et al., 1999). Thus, we used random-effects models in our analyses. In a single instance (AS > NS, experiment 1), we additionally considered results of a fixed-effects model, in which intersubject variability is not removed (see Results).

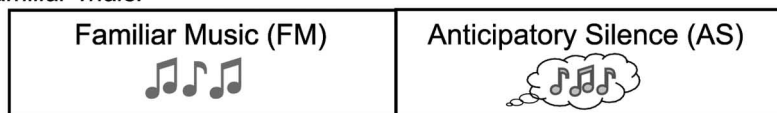
In all analyses, we selected GLM predictors corresponding to the four conditions: FM, UM, AS, and NS. In experiment 2, we excluded trials in which participants reported imagery “errors” from all contrasts of interest. These error trials were entered as predictors in GLMs but were discarded in subsequent contrasts (i.e., “predictors of no interest”). Thus,

A. Anticipatory Imagery



B. Experimental Design, fMRI Task

Familiar Trials:



Unfamiliar Trials:

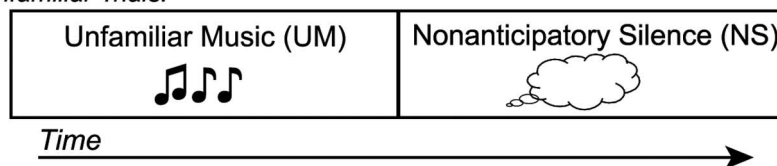


Figure 1. Experimental design. **A**, Schematic of anticipatory imagery. Hearing melody 1 triggers internal imagery for melody 2 before melody 2 is actually heard. **B**, Schematic of the behavioral task for both fMRI experiments. FM or UM was followed by silence. In the silence after FM, participants perceived imagery for the subsequent piece of music (AS). In the silence after UM, participants simply waited until the trial's end (NS).

FM and AS trials in which participants gave vividness ratings <4 on the 5-point scale were discarded from all analyses except the parametrically weighted GLM (see below); likewise, UM, NS, and silent baseline trials with ratings >1 were also discarded in experiment 2.

In whole-brain analyses, appropriate single-voxel map thresholds were chosen for random (E1: $t_{(8)} > 3.83, p < 0.005$; E2: $t_{(9)} < 3.69, p < 0.005$) and fixed-effects (E1 AS > NS: $t_{(2340)} > 2.81, p < 0.005$) analyses. Monte Carlo simulations were used to estimate the rate of false-positive responses, which we then used to obtain corrected cluster thresholds for these maps at $p(\text{corr}) < 0.05$ (Forman et al., 1995).

Parametrically weighted GLM analyses were conducted in experiment 2, to assess the relationships between the strength of the retrieved musical image indicated by behavioral ratings and BOLD–fMRI signal. Weights of the GLM predictors were adjusted to reflect vividness ratings 2–5 in a linear manner (–1.0, –0.33, 0.33, 1.0), whereas all other predictors were given a weight of 1. Thus, this analysis identified voxels having a BOLD–fMRI signal that was linearly correlated with vividness ratings on familiar trials.

ROI analyses were conducted in both experiments, using foci identified by relevant whole-brain GLM analyses. These random-effects GLM analyses were performed to explore subtler relationships between our conditions and the mean BOLD–fMRI signal within these predefined regions, including (1) the relative magnitude of response to FM and UM conditions in ROIs sensitive to music in general and (2) learning-related differences in AS ROIs across runs in experiment 2. Linear detrending and within-run normalization allow for signal comparisons between runs. Relevant bar graphs depict average percentage signal change from baseline within the relevant ROIs with SE across subjects.

Results

Experiment 1

In the first experiment, we measured neural responses associated with anticipatory imagery for highly familiar music. During the scan, participants heard the final 32 s of tracks from their favorite CD presented randomly (FM), followed by silence in which they would experience anticipatory imagery for the following track (AS). We also presented the final 32 s of unfamiliar tracks (UM),

which was again followed by a period of silence (NS) during which participants would not experience imagery. This design is outlined in Figure 1, and coordinates for significant clusters in all analyses can be found in Table 1. Preliminary data from experiment 1 have been presented previously in brief form (Rauschecker, 2005).

Perception of FM and UM

Listening to music, whether familiar or unfamiliar, was associated with robust activity in auditory areas across superior temporal cortex bilaterally, as well as the right amygdala (Fig. 2). This music-related activity was identified using a conjunction between those voxels significantly more active for the FM and UM conditions compared with baseline (NS). ROI analysis demonstrated that signal change was not different for FM and UM in these areas, suggesting that both the auditory cortex and, notably, the amygdala were equally involved when listening to FM and UM. The amygdala likely reflects emotion processing common to both FM and UM used in this study (Anderson and Phelps, 2001; Koelsch et al., 2006; Kleber et al., 2007).

Contrasting FM and UM conditions directly in the whole-head analysis, we saw that FM was associated with increased posterior cingulate cortex [PCC; Brodmann's area (BA) 23] and parahippocampal gyrus (PHG; BA 35/40) (Fig. 3A, red clusters) responses compared with UM. UM did not drive activity in any brain area significantly more than FM, as signified by the absence of blue clusters in Figure 3A. ROI analyses of signal change within the PCC and PHG showed that although FM drove activity in these areas significantly above baseline ($p < 0.0002$ for both), UM did not ($p = 0.27$ and 0.10 , respectively) (Fig. 3B). This indicates that these areas were exclusively involved in perception of highly familiar music.

Anticipatory imagery for familiar music

AS was associated with significant activation in a series of right frontal regions (Fig. 4). The superior frontal gyrus (SFG; BA10) was significantly active when using a random-effects analysis at our chosen thresholds [$p(\text{corr}) < 0.05$]. This is a finding novel to studies of musical imagery and likely reflects processes relevant to retrieval of highly familiar, richly complex naturalistic music. A fixed-effects analysis (which does not consider intersubject variability) at $p(\text{corr}) < 0.05$ revealed a more extensive pattern of activation, including the SFG cluster, presupplementary motor cortex (pre-SMA; BA6), dorsal premotor cortex (dPMC; BA6), and inferior frontal gyrus (IFG; BA44/45) adjacent to ventral premotor cortex (vPMC), all within the right hemisphere. Several of these latter areas are consistent with previous work on musical imagery (Zatorre et al., 1996; Halpern and Zatorre, 1999) and serve as a complement to the novel SFG result.

Experiment 2

Experiment 2 recreated the anticipatory imagery captured in experiment 1 within a more controlled setting. In a relatively short

Table 1. Talairach coordinates and cluster volumes (in mm³) in experiment 1

	<i>x, y, z</i>	Volume
FM and UM		
R auditory cortex	53, -15, 8.7	23,598
L auditory cortex	-50, -21, 8.9	28,691
R amygdala	25, -12, -13	441
FM > UM		
PCC	-2.4, -56, 21	2602
L PHG	-34, -29, -12	170
AS (imagery)		
R dPMC	51, -0.7, 47	818
R pre-SMA	4.8, 2.7, 61	298
R IFG	49, 18, 19	1583
R SFG	34, 48, -5	1080

R, Right; L, left.

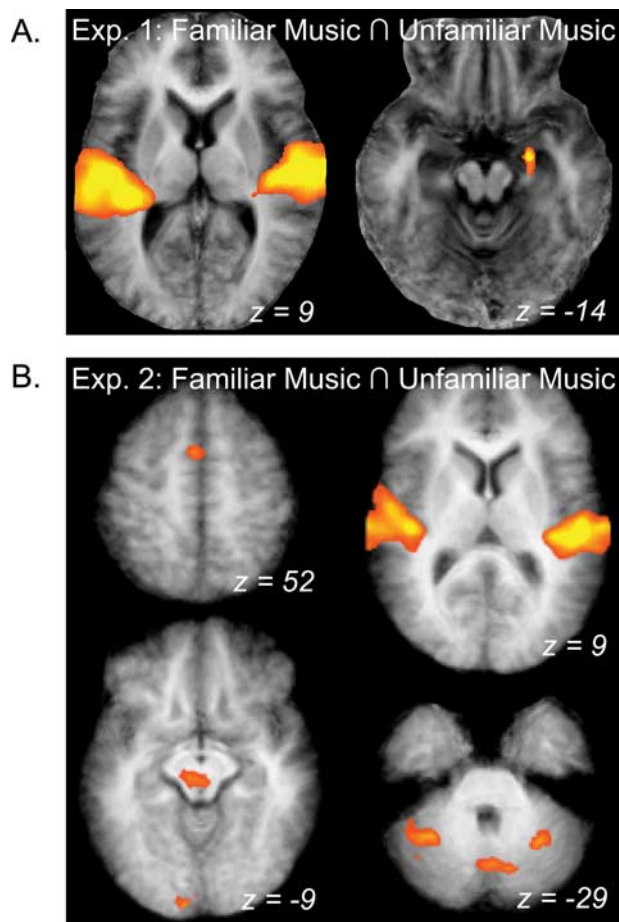


Figure 2. Areas responsive to FM and UM in experiments 1 and 2. **A**, Group activation maps demonstrating the results of a conjunction (\cap) between FM and UM conditions compared with baseline in a group-averaged anatomical image. Both familiar and unfamiliar naturalistic music elicited robust bilateral activity in large parts of the auditory cortex, as well as activity in the right amygdala in experiment 1 (Exp. 1). **B**, The same conjunction analysis reveals a wider pattern of activation associated with both familiar and unfamiliar simple melodies in experiment 2 (Exp. 2). Areas include bilateral auditory cortex (top right), presupplementary motor area (top left), midbrain and visual cortex (bottom left), and CB (bottom right). Results are shown in neurological coordinates (i.e., left is on the left).

45 min training session immediately before scanning, participants memorized seven pairs of short, simple melodies. BOLD-fMRI signal was then monitored while participants listened to and imagined melodies, similar to experiment 1 (Fig. 1). Experiment 2 included two identical and sequential runs, with a second short (<10 min) training phase occurring between these runs.

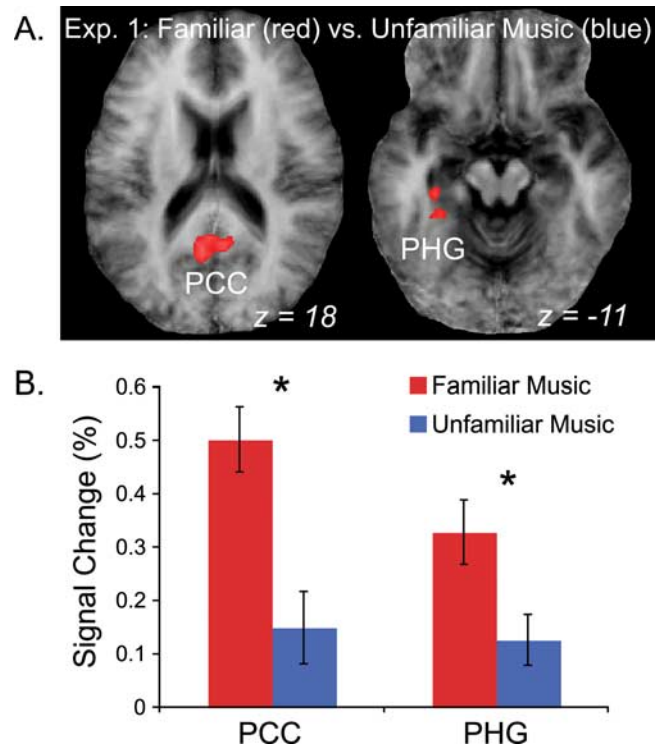


Figure 3. BOLD response to FM and UM in experiment 1 (Exp. 1). **A**, Group activation maps showing brain areas significantly more active for FM than UM, including PCC and PHG. **B**, Results of an ROI analysis of the clusters in **A**. Signal change above baseline is plotted along the vertical axis. The PCC and PHG were significantly more active for FM than baseline (red), but signal in these same ROIs was not significantly greater than baseline during the UM condition (blue). Error bars reflect SEM. Asterisks indicate significant differences between FM and UM used to define the clusters (* $p < 0.005$).

This design allowed us to assess differences in patterns of brain activation between moderately learned (run 1) and well learned (run 2) sequences. Coordinates for significant clusters in all experiment 2 analyses can be found in Table 2.

Behavioral results

Prescan training

After the initial training session and just before scanning, participants were tested on discrimination and recall accuracy (self-ratings) on trained melody pairs. In discriminating the second melody of trained pairs from distracter melodies with one-note deviations, participants achieved a high level of accuracy, making, on average, less than one error (mean accuracy, 93.65%; SD, 0.07). When asked to silently recall the second melody of trained pairs after being presented with the corresponding first melody, participants rated their images as being moderately accurate (mean rating, 6.03; SD, 1.73; 10-point rating scale) and moderately vivid (mean, 5.60; SD, 1.09; 10-point rating scale). Performance on this recall task indicated a moderate level of learning, which allowed room for improvement during the in-scan task and after the additional training between runs 1 and 2.

In-scan behavioral task

Vividness ratings on familiar trials (FM and AS) were significantly higher than those on unfamiliar trials (UM and NS) (significant main effect of condition; $F_{(3,27)} = 150.50$; $p < 0.001$). A significant run \times condition interaction indicates that this difference in vividness ratings increased between runs, as expected ($F_{(3,27)} = 7.76$; $p = 0.001$; run 1: familiar trials mean, 3.62; unfa-

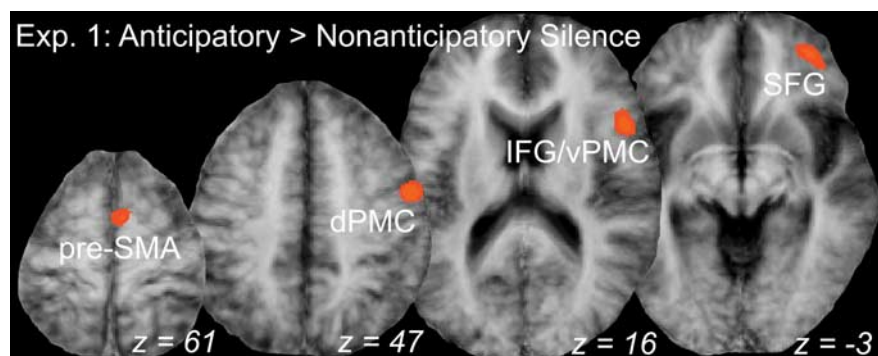


Figure 4. Brain areas responsive to anticipatory imagery in experiment 1 (Exp. 1). Group statistical maps of areas significantly more active for the AS condition compared with NS, including presupplementary motor area (pre-SMA), dPMC, IFG/vPMC, and SFG all in the right hemisphere. The results of the fixed-effects analysis are shown at $p(\text{corr}) < 0.05$.

Table 2. Talairach coordinates and cluster volumes (in mm³) in experiment 2

	<i>x, y, z</i>	Volume
FM and UM		
R auditory cortex	55, -13, 7	14,814
L auditory cortex	-50, -19, 7	11,416
pre-SMA	3, 5, 52	825
R cerebellum	31, -51, -28	1289
L cerebellum	-33, -51, -27	2235
CB vermis	0, -68, -24	2590
Visual cortex	-6, -75, 2	1846
	-15, -93, -4	730
Midbrain	-3, -24, -8	1081
FM > UM		
R pre-SMA/CMA	-9, 7, 47	658
L IPL	-39, -45, 40	2108
L IFG	-54, 5, 14	598
AS (imagery)		
SMA	-3, -7, 62	294
pre-SMA	0, 3, 54	122
PCC	-1, -32, 25	601
GP/Pu	-19, -5, 7	198
Thalamus	-15, -14, 7	94
	-3, -20, 16	983
L IPL	-35, -52, 36	59
R cerebellum	28, -53, -30	435
AS (weighted)		
R GP/Pu	23, 5, 2.8	911
L IFG	-47, 0, 12	457

R, Right; L, left.

miliar trials mean, 1.46; run 2: familiar trials mean, 4.01; unfamiliar trials mean, 1.36).

fMRI results

FM and UM

Listening to both familiar and unfamiliar melodies elicited robust bilateral activation of auditory areas along superior temporal cortex (Fig. 2), as in experiment 1. Music perception was also associated with increased activity in the pre-SMA (BA6), several foci along both hemispheres of the cerebellum (CB), the midbrain, and visual cortex. Comparing averaged signal change with these areas using an ROI analysis, no ROI was significantly more responsive to familiar or unfamiliar melodies, and signal was significantly greater than baseline (silence) for both conditions, suggesting that all areas were recruited in both conditions.

Contrasting FM and UM conditions directly across the entire brain, the pre-SMA and cingulate motor area (CMA; BA6/32),

left inferior parietal lobule (IPL; BA40), and left IFG or inferior vPMC (BA6/44) were significantly more active for familiar than unfamiliar musical sequences (Fig. 5A). Both FM and UM drove activity in most of these ROIs above baseline (Fig. 5B). However, signal in the IPL was not significantly greater than baseline for the UM condition, suggesting that the area was not involved in listening to unfamiliar music.

These results suggest that the inferior vPMC, pre-SMA, CMA, and CB are involved in the perception of both familiar and unfamiliar sequences. However, the frontal (vPMC, pre-SMA/CMA) and, in particular, the parietal (IPL) regions were

more involved in processing of familiar musical sequences, whereas the cerebellar activity was common to both familiar and unfamiliar melodies. Furthermore, the frontal areas (vPMC and pre-SMA/CMA) were similar to those activated in experiment 1 during imagery, which may indicate that these areas are involved in both perception and production of music.

Imagery for learned sequences

Comparing silence after learned sequences (AS) to silence after novel melodies (NS), imagery was associated with significant activity in the SMA proper (BA6), pre-SMA (BA6), PCC (BA23), left globus pallidus/putamen (GP/Pu), right lateral CB, left IPL (BA40), and thalamus (Fig. 6A,B). Unlike experiment 1, in which only frontal areas were recruited during the AS task, experiment 2 recruited a variety of both cortical and subcortical areas during imagery of musical sequences.

We performed an ROI analysis on these clusters to determine what areas are unaffected by learning (i.e., equally active in both runs) and what areas change during learning (i.e., differ in runs 1 and 2). In the first run, in which participants were moderately familiar with melody pairs, signal was significantly greater in the SMA proper (BA6) and GP/Pu than in run 2 (Fig. 6C). No ROI was more active in run 2 than in run 1. This indicates that the SMA proper and basal ganglia (BG) play a greater role in early or middle stages of learning. However, activity in all ROIs, including the SMA proper and GP/Pu, was significantly greater for the AS than the NS condition in both runs, indicating that these areas are important throughout the learning process. Interestingly, run 2 signal was only slightly greater for the AS than NS condition in GP/Pu ($p = 0.02$), suggesting that the BG may be most crucial to early stages of learning (Pasupathy and Miller, 2005).

We also identified areas in which fMRI signal correlated with in-scan behavioral vividness ratings using a parametrically weighted GLM. In this way, we were able to identify those areas with activity that increased with increasingly successful covert retrieval. Two areas demonstrated a positive linear correlation with participant vividness ratings during the AS condition: right GP/Pu and left inferior vPMC (BA 6/44) (Fig. 7A,B). Although the linear correlation between percentage signal change and behavioral rating was positive for both, the correlation was significant in both ROIs only in run 2, but not run 1 (run 1: GP/Pu, $r = 0.58$, $p = 0.08$; left vPMC, $r = 0.90$, $p = 0.11$; run 2: GP/Pu, $r = 0.96$, $p = 0.0006$; left vPMC, $r = 0.97$, $p = 0.003$). This indicates that signal in GP/Pu and vPMC more accurately reflects imagery processes in run 2 than in run 1, whereas the areas were recruited regardless of retrieval success in run 1. Additionally, because be-

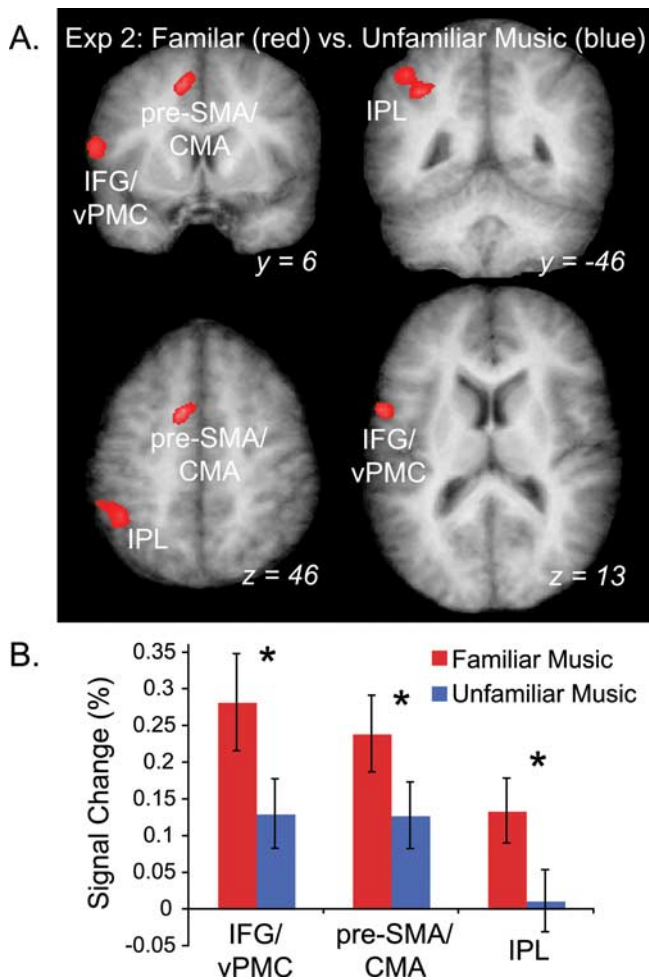


Figure 5. BOLD response to FM and UM in experiment 2 (Exp. 2). **A**, Group statistical maps of the FM versus UM contrast shown on a group-averaged anatomical image. Activation in three areas was greater for FM (red) than UM (blue) shown in coronal and horizontal sections, including IFG/vPMC, IPL, and presupplementary motor area/CMA (pre-SMA/CMA). No area was more active for UM than FM. **B**, Results of an ROI analysis on the clusters in **A** demonstrate that IPL signal was not significantly different during UM and baseline conditions. Error bars reflect SEM. Asterisks indicate significant differences between FM and UM used to define the clusters ($*p < 0.0001$).

havioral ratings were predictive of signal in these regions, we can be more confident in the validity of self-report measures gathered in these experiments. No area had signal negatively correlated with vividness ratings.

Discussion

The results presented here not only pinpoint brain areas involved in anticipatory retrieval of musical sequences but also explain how activity within this network changes throughout learning. Familiar music in experiment 1 had been heard by participants many more times, and over a longer period of time, than newly composed melodies in experiment 2. Thus, long familiar music from experiment 1 was likely to recruit areas involved in long-term encoding of music (including semantic and affective associations), whereas newly familiar music in experiment 2 was more likely to be processed in more “raw” form. Specifically, our data indicate that changes in frontal cortex and BG during learning, and perhaps ultimately in parietal and parahippocampal cortices as well, are key to developing increasingly complex brain representations of sequences.

Developing anticipatory imagery for musical sequences

While subjects imagined melodies after short-term explicit training in experiment 2, a variety of cortical (frontal and parietal) and subcortical (BG and CB) structures were involved. However, while experiencing imagery for highly familiar music after long-term incidental exposure in experiment 1, only frontal areas were recruited. This seems to indicate that at the end stages of musical learning, frontal cortex involvement dominates, whereas early and moderate stages require greater input from structures like the BG, which may “train” frontal networks (Pasupathy and Miller, 2005). Additionally, the CB was involved in anticipatory imagery in experiment 2, suggesting the CB’s role, like that of many so-called “motor” structures, reaches beyond overt production (Callan et al., 2006, 2007; Lahav et al., 2007).

Rostral prefrontal cortex

Imagery for highly familiar music in experiment 1 was associated with activity in a variety of frontal regions, including SFG, IFG/vPMC, dPMC, and pre-SMA. However, only the most rostral of these, SFG, was significant in the random-effects analysis, suggesting that SFG was more consistently involved during anticipatory imagery than the other frontal areas. Moreover, this rostral prefrontal area was not active when participants imagined moderately familiar music in experiment 2, further suggesting the area’s exclusive involvement in anticipatory imagery of highly familiar music. Lateral rostral prefrontal cortex is associated with episodic and working memory (Gilbert et al., 2006), including episodic memory for music (Platel et al., 2003) and prospective memory (Burgess et al., 2003; Sakai and Passingham, 2003; Gilbert et al., 2006). A caudal-to-rostral hierarchy within prefrontal cortex has been proposed (Christoff and Gabrieli, 2000; Sakai and Passingham, 2003), with rostral areas coordinating activity in more caudal modality-specific areas (Sakai and Passingham, 2003; Rowe et al., 2007). Our results indicate rostral prefrontal cortex could very likely be involved in cued recall of well learned, highly familiar music (i.e., anticipatory imagery) as well, perhaps through interactions with more caudal prefrontal and premotor regions.

Inferior frontal/vPMC

Our data indicate that inferior vPMC is involved in anticipatory cued recall of entire melodies in later stages of learning novel melody pairs (experiment 2), as well as during retrieval of highly familiar melody associations (experiment 1). Similar vPMC regions, as well as adjacent areas of the IFG located more rostrally, are involved in temporal integration of sound (Griffiths et al., 2000). This includes tasks requiring application of “syntactic” rules in music (Janata et al., 2002; Tillmann et al., 2006; Leino et al., 2007) and language (Friederici et al., 2003; Opitz and Friederici, 2007), and a role in congenital amusia (Hyde et al., 2007). vPMC also responds similarly to motor, visual, and imagery tasks, indicating that it performs these integrative computations in a domain-general manner (Schubotz and von Cramon, 2004; Schubotz, 2007). Inferior vPMC involvement during anticipatory imagery in our studies supports theories that this region is amenable to predictive, anticipatory processing of sequences.

BG and prefrontal cortex

The BG were most active during the earliest stage of learning in these experiments, the first run of experiment 2, and significantly decreased in activation in the second run of experiment 2 to a level barely above baseline. Interestingly, there is evidence that the BG are recruited early when learning stimulus associations but are less involved as learning progresses (Pasupathy and

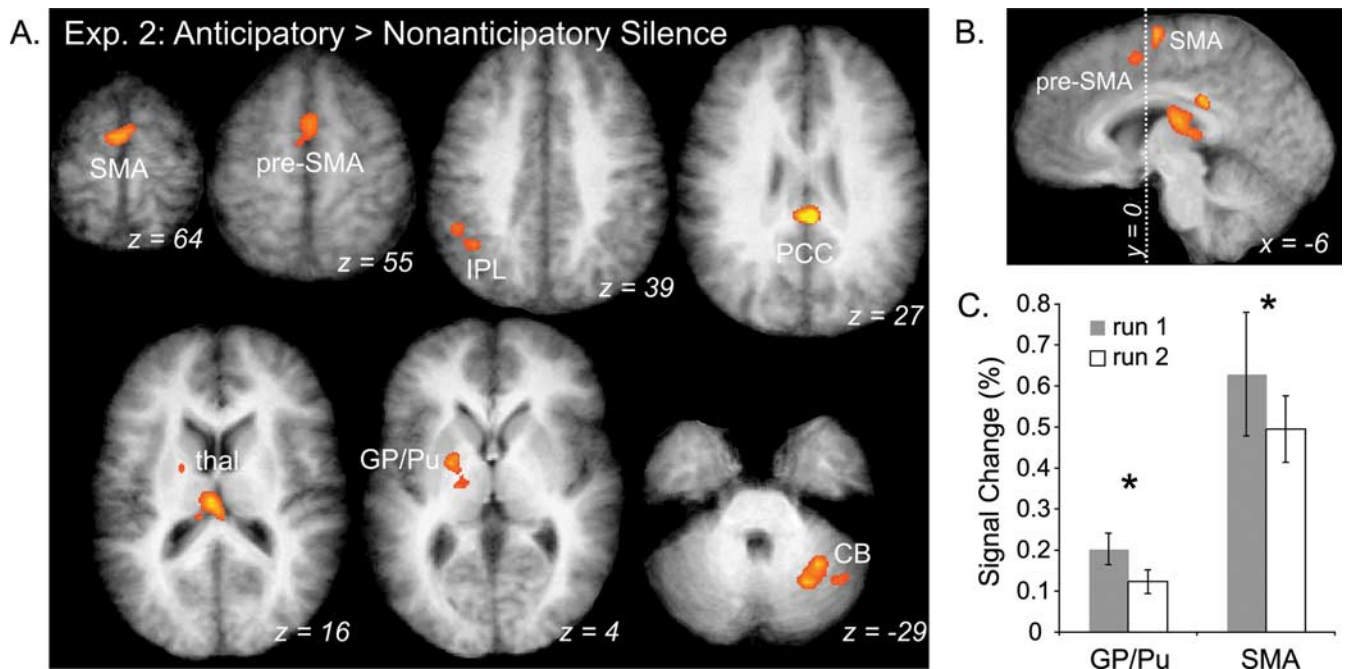


Figure 6. Activation maps of imagery during AS condition show learning effects in experiment 2 (Exp. 2). **A**, Significant activation associated with anticipatory imagery (AS > NS). Areas include CB, GP/Pu, thalamus, PCC, presupplementary motor area (pre-SMA), and SMA proper. **B**, Sagittal view of medial frontal activation. Dotted line indicating Talairach coordinate axis, $y = 0$, separates pre-SMA and SMA proper. **C**, ROI analysis reveals percentage signal change differences in the AS conditions compared between run 1 (shaded) and run 2 (white) ($*p < 0.05$). ROIs were defined by analysis shown in **A**. Error bars indicate SEM.

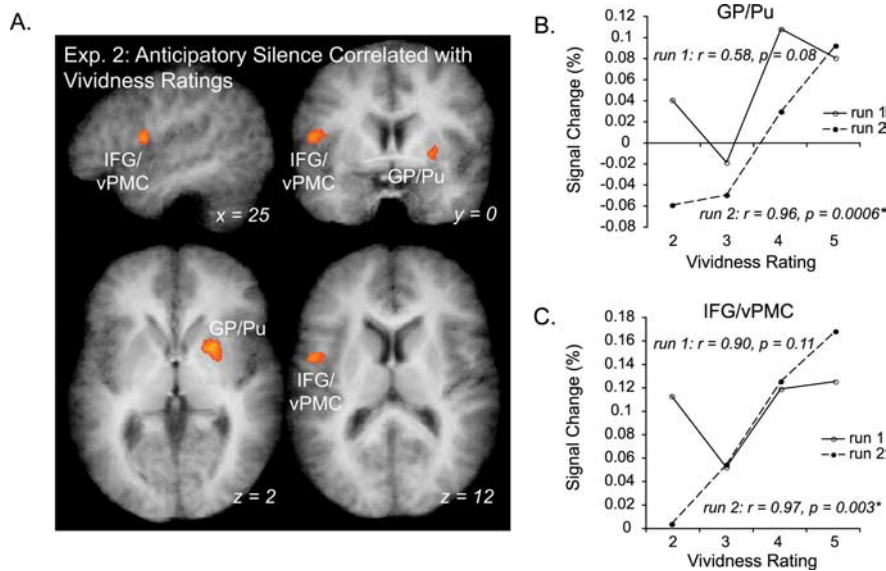


Figure 7. Areas for which BOLD signal is correlated with vividness rating in the familiar silence condition (experiment 2). **A**, Group activation maps reflect significant foci (GP/Pu, IFG/vPMC) resulting from the parametrically weighted GLM analysis, with model predictors adjusted to reflect participants' imagery vividness ratings during scanning. **B, C**, Line plots show correlations between percentage signal change and vividness ratings for ROIs resulting from the parametrically weighted GLM analysis in **A** (GP/Pu and IFG/vPMC, respectively). Solid lines correspond to results from the first run, and dashed lines correspond to data from run 2, after additional training. Exp. 2, Experiment 2.

Miller, 2005; Poldrack et al., 2005). This has led to the hypothesis that the BG may function to prepare or train prefrontal areas (Pasupathy and Miller, 2005). In songbirds, anterior forebrain and BG nuclei are critical in sensorimotor song learning during development (Brainard and Doupe, 2002) and modulation of song production in adults (Kao et al., 2005). Indeed, the BG and inferior frontal cortex were the only two areas in experiment 2

that demonstrated significant correlations between behavioral ratings and BOLD-fMRI signal, further suggesting the importance of the relationship between these two regions, especially when retrieving newly learned associations.

Supplementary motor areas

Signal in SMA proper during anticipatory imagery was inversely proportional to the degree of learning in our experiments. Pre-SMA, in contrast, was activated during both experiments (although subthreshold activation in the random-effects analysis in experiment 1 suggests less consistent recruitment). SMA proper is more closely related to primary motor areas and movement execution (Hikosaka et al., 1996; Weille et al., 2001; Picard and Strick, 2003) and is typically active throughout sequence execution (Matsuzaka et al., 1992). Pre-SMA, in contrast, is transiently active at sequence initiation (Matsuzaka et al., 1992) and is associated with more complex motor action plans (Alario et al., 2006), including chunks of sequence items (Sakai et al., 1999; Kennerley et al., 2004). Here, pre-SMA was involved in retrieval of more complex learned representations (i.e., musical phrases or "chunks" of sequence material), whereas SMA was recruited during earlier stages that presumably rely more on individual note-to-note associations. This assessment is compatible with previous music imagery studies that report pre-SMA involvement in imagery for familiar melodies (Halpern and Zatorre, 1999) and SMA involvement in imagery for single

sounds and sequence repetition (Halpern and Zatorre, 1999; Halpern et al., 2004), neither of which is likely to rely on “chunking” of multiple sequence items.

Familiar music recruits structures involved in memory

Brain structures involved in perception of familiar music, but not unfamiliar music, included the PHG and PCC in experiment 1 and an IPL subregion in experiment 2. The hippocampus and surrounding medial temporal lobe have long been linked to memory processes (Schacter et al., 1998; Squire et al., 2004), including memory for music (Watanabe et al., 2008) and consolidation of learned motor sequences (Albouy et al., 2008). PCC and adjacent medial parietal cortex share connections with medial temporal regions (Burwell and Amaral, 1998; Kahn et al., 2008), and damage to PCC can result in memory deficits (Valenstein et al., 1987; Osawa et al., 2008), suggesting that the area is more directly involved in memory retrieval. Thus, the PHG and PCC are likely to be involved in recognizing familiar music after consolidation of that music into long-term memory (i.e., experiment 1).

Lateral parietal cortex is also connected with the medial temporal lobe (Burwell and Amaral, 1998; Kahn et al., 2008) and is routinely recruited during memory tasks (Wagner et al., 2005; Cabeza et al., 2008). Lateral parietal damage, however, does not seem to confer measurable stimulus recognition deficits (Cabeza et al., 2008; Haramati et al., 2008), which has led to the idea that the IPL may be involved in the mediation of memory effects on behavior (Wagner et al., 2005) or attention to cued memories (Cabeza et al., 2008). In our study, the IPL was associated with listening to newly learned music in experiment 2, but was also recruited during anticipatory imagery in experiment 2, suggesting the IPL may mediate both recall and recognition of newly learned musical sequences.

Conclusions

Our results indicate that learning sound sequences recruits similar brain structures as motor sequence learning (Hikosaka et al., 2002) and, hence, similar principles might be at work in both domains. Moreover, the very existence of anticipatory imagery suggests that retrieving stored sequences of any kind involves predictive readout of upcoming information before the actual sensorimotor event. This prospective signal may be related to the “efference copy” proposed in previous models (von Holst and Mittelstaedt, 1950; Troyer and Doupe, 2000), including so-called emulator (or “forward”) models (Grush, 2004; Callan et al., 2007; Ghasia et al., 2008). As discussed above, prefrontal cortex is a prime location for developing complex stimulus associations required to process sequences of events. This is confirmed both by strong frontal involvement during our task, as well as the caudal-to-rostral progression of frontal involvement during different stages of learning, possibly reflecting ascending levels in the processing hierarchy. Other structures are also likely to be involved in these putative internal models at various stages, including ventral premotor and posterior parietal cortices, BG, and CB. Future research confirming the temporal order of involvement of these structures is needed (Sack et al., 2008), and studies of auditory-perceptual or audiomotor learning in nonhuman primates could serve as a bridge between electrophysiology in songbirds (Dave and Margoliash, 2000; Troyer and Doupe, 2000) and human imaging.

References

Alario FX, Chainay H, Lehericy S, Cohen L (2006) The role of the supplementary motor area (SMA) in word production. *Brain Res* 1076:129–143.

- Albouy G, Sterpenich V, Baeteu E, Vandewalle G, Desseilles M, Dang-Vu T, Darsaud A, Ruby P, Luppi PH, Degueldre C, Peigneux P, Luxen A, Maquet P (2008) Both the hippocampus and striatum are involved in consolidation of motor sequence memory. *Neuron* 58:261–272.
- Anderson AK, Phelps EA (2001) Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature* 411:305–309.
- Brainard MS, Doupe AJ (2002) What songbirds teach us about learning. *Nature* 417:351–358.
- Burgess PW, Scott SK, Frith CD (2003) The role of the rostral frontal cortex (area 10) in prospective memory: a lateral versus medial dissociation. *Neuropsychologia* 41:906–918.
- Burwell RD, Amaral DG (1998) Cortical afferents of the perirhinal, postirhinal, and entorhinal cortices of the rat. *J Comp Neurol* 398:179–205.
- Cabeza R, Ciaramelli E, Olson IR, Moscovitch M (2008) The parietal cortex and episodic memory: an attentional account. *Nat Rev Neurosci* 9:613–625.
- Callan DE, Tsytsev V, Hanakawa T, Callan AM, Katsuhara M, Fukuyama H, Turner R (2006) Song and speech: brain regions involved with perception and covert production. *Neuroimage* 31:1327–1342.
- Callan DE, Kawato M, Parsons L, Turner R (2007) Speech and song: the role of the cerebellum. *Cerebellum* 6:321–327.
- Christoff K, Gabrieli JDE (2000) Frontopolar cortex and human cognition: evidence for a rostrocaudal hierarchical organization within the human prefrontal cortex. *Psychobiology* 28:168–186.
- Dave AS, Margoliash D (2000) Song replay during sleep and computational rules for sensorimotor vocal learning. *Science* 290:812–816.
- Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC (1995) Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magn Reson Med* 33:636–647.
- Friederici AD, Ruschmeyer SA, Hahne A, Fiebach CJ (2003) The role of left inferior frontal and superior temporal cortex in sentence comprehension: localizing syntactic and semantic processes. *Cereb Cortex* 13:170–177.
- Friston KJ, Holmes AP, Poline JB, Grasby PJ, Williams SC, Frackowiak RS, Turner R (1995) Analysis of fMRI time-series revisited. *Neuroimage* 2:45–53.
- Ghasia FF, Meng H, Angelaki DE (2008) Neural correlates of forward and inverse models for eye movements: evidence from three-dimensional kinematics. *J Neurosci* 28:5082–5087.
- Gilbert SJ, Spengler S, Simons JS, Steele JD, Lawrie SM, Frith CD, Burgess PW (2006) Functional specialization within rostral prefrontal cortex (area 10): a meta-analysis. *J Cogn Neurosci* 18:932–948.
- Griffiths TD, Penhune V, Peretz I, Dean JL, Patterson RD, Green GG (2000) Frontal processing and auditory perception. *Neuroreport* 11:919–922.
- 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.
- Halpern AR, Zatorre RJ (1999) When that tune runs through your head: a PET investigation of auditory imagery for familiar melodies. *Cereb Cortex* 9:697–704.
- Halpern AR, Zatorre RJ, Bouffard M, Johnson JA (2004) Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia* 42:1281–1292.
- Haramati S, Soroker N, Dudai Y, Levy DA (2008) The posterior parietal cortex in recognition memory: a neuropsychological study. *Neuropsychologia* 46:1756–1766.
- Hikosaka O, Sakai K, Miyauchi S, Takino R, Sasaki Y, Putz B (1996) Activation of human presupplementary motor area in learning of sequential procedures: a functional MRI study. *J Neurophysiol* 76:617–621.
- Hikosaka O, Nakamura K, Sakai K, Nakahara H (2002) Central mechanisms of motor skill learning. *Curr Opin Neurobiol* 12:217–222.
- Hyde KL, Lerch JP, Zatorre RJ, Griffiths TD, Evans AC, Peretz I (2007) Cortical thickness in congenital amusia: when less is better than more. *J Neurosci* 27:13028–13032.
- Janata P, Grafton ST (2003) Swinging in the brain: shared neural substrates for behaviors related to sequencing and music. *Nat Neurosci* 6:682–687.
- Janata P, Birk JL, Van Horn JD, Leman M, Tillmann B, Bharucha JJ (2002) The cortical topography of tonal structures underlying Western music. *Science* 298:2167–2170.

- Kahn I, Andrews-Hanna JR, Vincent JL, Snyder AZ, Buckner RL (2008) Distinct cortical anatomy linked to subregions of the medial temporal lobe revealed by intrinsic functional connectivity. *J Neurophysiol* 100:129–139.
- Kao MH, Doupe AJ, Brainard MS (2005) Contributions of an avian basal ganglia-forebrain circuit to real-time modulation of song. *Nature* 433:638–643.
- Kennerley SW, Sakai K, Rushworth MF (2004) Organization of action sequences and the role of the pre-SMA. *J Neurophysiol* 91:978–993.
- Kleber B, Birbaumer N, Veit R, Trevorrow T, Lotze M (2007) Overt and imagined singing of an Italian aria. *Neuroimage* 36:889–900.
- Koelsch S, Fritz T, V Cramon DY, Muller K, Friederici AD (2006) Investigating emotion with music: an fMRI study. *Hum Brain Mapp* 27:239–250.
- 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.
- Leino S, Brattico E, Tervaniemi M, Vuust P (2007) Representation of harmony rules in the human brain: further evidence from event-related potentials. *Brain Res* 1142:169–177.
- Margoliash D, Fortune ES (1992) Temporal and harmonic combination-sensitive neurons in the zebra finch's HVC. *J Neurosci* 12:4309–4326.
- Matsuzaka Y, Aizawa H, Tanji J (1992) A motor area rostral to the supplementary motor area (presupplementary motor area) in the monkey: neuronal activity during a learned motor task. *J Neurophysiol* 68:653–662.
- Opitz B, Friederici AD (2007) Neural basis of processing sequential and hierarchical syntactic structures. *Hum Brain Mapp* 28:585–592.
- Osawa A, Maeshima S, Kunishio K (2008) Topographic disorientation and amnesia due to cerebral hemorrhage in the left retrosplenial region. *Eur Neurol* 59:79–82.
- Pasupathy A, Miller EK (2005) Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature* 433:873–876.
- Perry DW, Zatorre RJ, Petrides M, Alivisatos B, Meyer E, Evans AC (1999) Localization of cerebral activity during simple singing. *Neuroreport* 10:3979–3984.
- Petersson KM, Nichols TE, Poline JB, Holmes AP (1999) Statistical limitations in functional neuroimaging. I. Non-inferential methods and statistical models. *Philos Trans R Soc Lond B Biol Sci* 354:1239–1260.
- Picard N, Strick PL (2003) Activation of the supplementary motor area (SMA) during performance of visually guided movements. *Cereb Cortex* 13:977–986.
- Plailly J, Tillmann B, Royet JP (2007) The feeling of familiarity of music and odors: the same neural signature? *Cereb Cortex* 17:2650–2658.
- Platel H, Baron JC, Desgranges B, Bernard F, Eustache F (2003) Semantic and episodic memory of music are subserved by distinct neural networks. *Neuroimage* 20:244–256.
- Poldrack RA, Sabb FW, Foerde K, Tom SM, Asarnow RF, Bookheimer SY, Knowlton BJ (2005) The neural correlates of motor skill automaticity. *J Neurosci* 25:5356–5364.
- Rauschecker JP (2005) Neural encoding and retrieval of sound sequences. *Ann N Y Acad Sci* 1060:125–135.
- Rauschecker JP, Tian B, Hauser M (1995) Processing of complex sounds in the macaque nonprimary auditory cortex. *Science* 268:111–114.
- Rowe JB, Sakai K, Lund TE, Ramsøy T, Christensen MS, Baare WF, Paulson OB, Passingham RE (2007) Is the prefrontal cortex necessary for establishing cognitive sets? *J Neurosci* 27:13303–13310.
- Sack AT, Jacobs C, De Martino F, Staeren N, Goebel R, Formisano E (2008) Dynamic premotor-to-parietal interactions during spatial imagery. *J Neurosci* 28:8417–8429.
- Sakai K, Passingham RE (2003) Prefrontal interactions reflect future task operations. *Nat Neurosci* 6:75–81.
- Sakai K, Hikosaka O, Miyauchi S, Sasaki Y, Fujimaki N, Putz B (1999) Pre-supplementary motor area activation during sequence learning reflects visuo-motor association. *J Neurosci* 19:RC1(1–6).
- Schacter DL, Buckner RL, Koutstaal W (1998) Memory, consciousness and neuroimaging. *Philos Trans R Soc Lond B Biol Sci* 353:1861–1878.
- Schubotz RI (2007) Prediction of external events with our motor system: towards a new framework. *Trends Cogn Sci* 11:211–218.
- Schubotz RI, von Cramon DY (2003) Functional-anatomical concepts of human premotor cortex: evidence from fMRI and PET studies. *Neuroimage* 20 [Suppl 1]:S120–S131.
- Schubotz RI, von Cramon DY (2004) Sequences of abstract nonbiological stimuli share ventral premotor cortex with action observation and imagery. *J Neurosci* 24:5467–5474.
- Sergent J, Zuck E, Terriah S, MacDonald B (1992) Distributed neural network underlying musical sight-reading and keyboard performance. *Science* 257:106–109.
- Squire LR, Stark CE, Clark RE (2004) The medial temporal lobe. *Annu Rev Neurosci* 27:279–306.
- Suga N, O'Neill WE, Manabe T (1978) Cortical neurons sensitive to combinations of information-bearing elements of biosonar signals in the mustache bat. *Science* 200:778–781.
- Talairach J, Tournoux P (1988) Co-planar stereotaxic atlas of the human brain. Stuttgart: Thieme.
- Tillmann B, Koelsch S, Escoffier N, Bigand E, Lalitte P, Friederici AD, von Cramon DY (2006) Cognitive priming in sung and instrumental music: activation of inferior frontal cortex. *Neuroimage* 31:1771–1782.
- Troyer TW, Doupe AJ (2000) An associational model of birdsong sensorimotor learning I. Efference copy and the learning of song syllables. *J Neurophysiol* 84:1204–1223.
- Valenstein E, Bowers D, Verfaellie M, Heilman KM, Day A, Watson RT (1987) Retrosplenial amnesia. *Brain* 110:1631–1646.
- von Holst E, Mittelstaedt H (1950) Das Reafferenzprinzip: Wechselwirkungen zwischen Zentralnervensystem und Peripherie. *Naturwissenschaften* 37:464–476.
- Wagner AD, Shannon BJ, Kahn I, Buckner RL (2005) Parietal lobe contributions to episodic memory retrieval. *Trends Cogn Sci* 9:445–453.
- Watanabe T, Yagishita S, Kikyo H (2008) Memory of music: roles of right hippocampus and left inferior frontal gyrus. *Neuroimage* 39:483–491.
- Weilke F, Spiegel S, Boecker H, von Einsiedel HG, Conrad B, Schwaiger M, Erhard P (2001) Time-resolved fMRI of activation patterns in M1 and SMA during complex voluntary movement. *J Neurophysiol* 85:1858–1863.
- Zatorre RJ, Halpern AR, Perry DW, Meyer E, Evans AC (1996) Hearing in the mind's ear: a PET investigation of musical imagery and perception. *J Cogn Neurosci* 8:29–46.
- 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.