

Bucknell University

Bucknell Digital Commons

Faculty Journal Articles

Faculty Scholarship

2012

Neuronal Correlates of Perception, Imagery, and Memory for Familiar Tunes

Sibylle C. Herholz

Andrea R. Halpern

Bucknell University, ahalpern@bucknell.edu

Robert J. Zatorre

Follow this and additional works at: https://digitalcommons.bucknell.edu/fac_journ



Part of the [Psychology Commons](#)

Recommended Citation

Herholz, Sibylle C.; Halpern, Andrea R.; and Zatorre, Robert J.. "Neuronal Correlates of Perception, Imagery, and Memory for Familiar Tunes." *Journal of Cognitive Neuroscience* (2012) : 1382-1397.

This Article is brought to you for free and open access by the Faculty Scholarship at Bucknell Digital Commons. It has been accepted for inclusion in Faculty Journal Articles by an authorized administrator of Bucknell Digital Commons. For more information, please contact dcadmin@bucknell.edu.

Neuronal Correlates of Perception, Imagery, and Memory for Familiar Tunes

Sibylle C. Herholz¹, Andrea R. Halpern², and Robert J. Zatorre¹

Abstract

■ We used fMRI to investigate the neuronal correlates of encoding and recognizing heard and imagined melodies. Ten participants were shown lyrics of familiar verbal tunes; they either heard the tune along with the lyrics, or they had to imagine it. In a subsequent surprise recognition test, they had to identify the titles of tunes that they had heard or imagined earlier. The functional data showed substantial overlap during melody perception and imagery, including secondary auditory areas. During imagery compared with perception, an extended network including pFC, SMA, intraparietal sulcus, and cerebellum showed increased activity, in line with the increased processing demands of imagery. Functional connectivity of anterior right temporal

cortex with frontal areas was increased during imagery compared with perception, indicating that these areas form an imagery-related network. Activity in right superior temporal gyrus and pFC was correlated with the subjective rating of imagery vividness. Similar to the encoding phase, the recognition task recruited overlapping areas, including inferior frontal cortex associated with memory retrieval, as well as left middle temporal gyrus. The results present new evidence for the cortical network underlying goal-directed auditory imagery, with a prominent role of the right pFC both for the subjective impression of imagery vividness and for on-line mental monitoring of imagery-related activity in auditory areas. ■

INTRODUCTION

Imagine putting on a recording of a favorite song or musical piece. Even if one is not a trained musician, it is possible to come up with a mental auditory image that resembles the real experience of hearing that song even before it starts playing. Auditory imagery refers to this aspect of auditory cognition in which auditory information is internally generated and processed in the absence of real sound perception. It can be surprisingly accurate, reinstating many aspects of the real stimulus in the mind (Janata & Paroo, 2006; Halpern, Zatorre, Bouffard, & Johnson, 2004; Crowder, 1989; Halpern, 1989). Previous neuroimaging studies have shown that during voluntary auditory imagery of music, secondary auditory cortex and association areas are active (Halpern et al., 2004; Halpern & Zatorre, 1999; Zatorre, Halpern, Perry, Meyer, & Evans, 1996). Support for the involvement of secondary auditory cortices during musical imagery also comes from magnetoencephalography and EEG studies (Herholz, Lappe, Knief, & Pantev, 2008; Schürmann, Raji, Fujiki, & Hari, 2002; Janata, 2001). However, little is known about the topographic specificity of cortical areas that are involved in auditory imagery, in part because very few studies have directly compared areas active in real versus imagined sound. Therefore, a first aim of this study was to directly compare activity related to imagery and perception of real melodies in otherwise identical encoding tasks.

Mental imagery not only relies on auditory cortex activity but is also associated with increased activity in several other regions including the SMA, frontal, and parietal areas (Halpern & Zatorre, 1999; Zatorre et al., 1996). The pFC, together with motor and premotor areas, is also involved in anticipatory imagery, when participants have learned to anticipate successive presentations of familiar songs (Leaver, Van Lare, Zielinski, Halpern, & Rauschecker, 2009). Whereas activity in secondary auditory areas is assumed to underlie the subjective impression of sound during auditory imagery (Halpern et al., 2004; Schürmann et al., 2002; Halpern & Zatorre, 1999; Zatorre et al., 1996), premotor areas and SMA seem to be involved in subvocalization or mental singing (Halpern & Zatorre, 1999; Zatorre et al., 1996; but see also Halpern et al., 2004), and frontal areas are assumed to support memory retrieval of the sounds in case of familiar material, and working memory in image generation (Halpern & Zatorre, 1999).

On the basis of these studies, we have a good picture of which areas are involved in auditory imagery, but how do they interact? So far, to our knowledge, no one has investigated the functional connections within the cortical network during auditory imagery. During the last decade, neuroimaging has moved from localization of specific areas involved in a task to the investigation of interactions between cortical and subcortical areas. For mental imagery in the visual domain, Mechelli, Price, Friston, and Ishai (2004) have shown that the top-down connectivity of prefrontal and visual areas is increased during imagery

¹McGill University, ²Bucknell University

compared with perception (Mechelli et al., 2004). This top-down influence of frontal areas on the specific sensory areas involved in imagery seems plausible for auditory imagery as well, as both pFC and secondary auditory cortex are involved in auditory imagery. Therefore, another principal aim of this study was to investigate the functional network underlying mental imagery in the auditory domain. In line with the findings from vision, we hypothesized that auditory imagery should involve stronger functional connections of auditory with frontal and prefrontal areas compared with networks active during perception of those same melodies.

Whereas the subjective experience in voluntary imagery is not normally confused for real perception, mental imagery can nonetheless be a very vivid experience. During auditory imagery, most aspects of perception, including pitch (Halpern, 1989), tempo (Halpern, 1988) and timbre (Halpern et al., 2004; Crowder, 1989), seem to be preserved. Two recent studies have successfully identified brain correlates of imagery vividness on a trial-by-trial basis in the auditory cortex (Daselaar, Porat, Huijbers, & Pennartz, 2010), in the left inferior frontal gyrus/ventral premotor cortex, and in the right globus pallidus/putamen (Leaver et al., 2009). However, vividness of imagery is not only a state that can vary across trials; the ability to vividly imagine auditory information also appears to be a stable trait that varies among people (White, Ashton, & Brown, 1977). Let us revisit our example of the music recording: One person's mental impression might be rather vague and pale, maybe just including the melody, whereas someone else's might be vivid, lifelike, and rich in details such as the timbres of the instrumentation, the emotional expressions or the dynamics. Whereas most previous studies on imagery in the auditory domain have focused on group averages and have treated differences in imagery vividness as noise, data from other fields suggest that individual differences can be an important explanatory variable in neuronal correlates of both basic sensory processes and higher-order cognition (Kanai & Rees, 2011).

In the visual domain, Cui and colleagues showed that individual differences in vividness of visual imagery can be related to brain activity in visual cortex during visual imagery (Cui, Jeter, Yang, Montague, & Eagleman, 2007). Similarly, visual imagery vividness scores predict the amount of suppression of auditory cortex activity during visual imagery (Amedi, Malach, & Pascual-Leone, 2005). On the basis of such findings from vision research, we hypothesized that general ability for vivid imagery influences neuronal correlates during mental imagery also in the auditory domain. In a recent study, expert musicians' self-reported vividness of auditory imagery correlated with activity in the right intraparietal sulcus (IPS) during the mental reversal of melodies, a task that heavily relies on mental imagery and mental manipulation of sounds (Zatorre, Halpern, & Bouffard, 2010). Although these results provided first evidence that individual differences

in imagery vividness modulate neuronal correlates in related tasks in the auditory domain, the highly trained group of participants and the difficult and complex mental reversal task made extrapolation of these findings to auditory imagery in general difficult. Therefore, in this study, our third main goal was to determine if imagery-related activity during a simple task of imagining familiar melodies is modulated by a person's overall vividness of auditory imagery, across a range of musical training.

Auditory cortices are not only important for imagery but also for retrieval of auditory information from memory (Wheeler, Petersen, & Buckner, 2000), and there seems to be some overlap between the two systems (Zatorre & Halpern, 2005; Peretz & Coltheart, 2003). A recent study showed directly that neural correlates of imagery and recognition memory overlap in secondary auditory areas (Huijbers, Pennartz, Rubin, & Daselaar, 2011). However, in this study imagery of melodies was only compared with retrieval of perceived and not imagined melodies. Another interesting question is how we differentiate episodic memories of real perception from those of imagery and how the retrieval of this information can be related to brain activity during encoding. The question of how episodic memories of real and imagined events differ is not only interesting for theories of memory encoding and retrieval but also for more practical applications, for example, in legal or clinical context (Loftus, 2005), as mental imagery can result in false autobiographical memories (Hyman & Pentland, 1996). Very vivid imagery is potentially detrimental to the task of remembering actually perceived or imagined events. A vivid image might generate a memory representation with strong perceptual qualities to it, thereby increasing the similarity between heard and imagined tunes. We were interested in looking at a type of false memory in which a person who has only imagined a familiar song later claims he or she actually heard it. This raises the interesting question of whether activation of imagery-related brain areas will be correlated not with correct performance, but in fact incorrect performance. A study in the visual domain (Gonsalves et al., 2004) showed such effects for false memories of having seen photographs, but no one has extended this finding to the auditory domain. We expected that strong auditory imagery vividness should lead to more source memory errors and that brain correlates for items that were later correctly or incorrectly remembered as imagined or heard would differ, analogous to findings in the visual domain (Gonsalves et al., 2004).

In summary, the aims of the current study were (1) to identify overlapping and unique areas related to imagery and perception of familiar tunes, (2) to identify a functional network connecting auditory areas to other cortical areas involved in mental imagery of familiar tunes, (3) to investigate how the subjective vividness of imagery relates to cortical activity during musical imagery, and (4) to investigate which areas are involved in the recognition of previously imagined or perceived musical stimuli.

METHODS

Participants

Ten healthy, right-handed participants (five women, aged 20–38 years, mean = 27 years) without any neurological or psychiatric conditions and with normal hearing capacities participated in the study. All participants had grown up in Canada or in the United States and had English as their mother tongue, which ensured that they were familiar with the tunes used in the experiment. Participants gave written consent before participation and received monetary compensation for their time and inconvenience. The research ethics board of the Montreal Neurological Institute, McGill University, approved all procedures. Before the start of the experiment, participants completed questionnaires about demographic data, handedness, and musical training background. We used the Bucknell Auditory Imagery Scale (BAIS) to assess the vividness of auditory images and the flexibility/control of imagery, that is, the ability to change or manipulate an auditory image at will. This questionnaire has previously been used by our laboratories as a measure of the general vividness of auditory imagery and mental control of auditory imagery (Zatorre et al., 2010). The BAIS assesses both aspects of imagery in two subscales encompassing 14 items each. In each item, an auditory scenario is described, and the subject's task is to imagine the scenario and rate the vividness of their auditory imagery on a scale from 1 to 7. As we were mainly interested in the vividness rather than the control of auditory imagery, we used the vividness subscore in our analyses. Participants for the main study were selected so that they represented a range of both vividness scores (range = 4–6.64, mean = 5.6) and musical experience (range = 0–27 years), with a low correlation between them ($r = .254$). This enabled us to investigate the effect of imagery vividness without the possible confounding effect of musical training on the data.

Stimuli

Short samples of 90 familiar songs were used as stimuli. The songs were selected from a set of 160 songs based on their familiarity values in a piloting phase with a different group of 10 participants who were comparable in age and cultural and musical background with the participants of the main study. All selected songs had mean familiarity ratings of >5 on a scale of 1–7 (7 = *very familiar*). Songs were divided into three groups, with an approximately equal number of songs of different categories (e.g., children's tunes, Christmas carols, pop songs, etc.) and familiarity values in each group. Assignment of the songs to each of the conditions in the experiment (Imagine, Listen, New) was balanced across participants. Thus, although each individual did not encounter all of the melodies in the first phase of the experiment (only heard and imagined melodies, not the “new” melodies that were presented as lures during the recognition test), on average

all of the melodies featured equally often in each of the conditions across participants.

Excerpts were either the beginning of the chorus or the verse of the songs, whichever was better known. Lengths of the excerpts varied between 4 and 8 sec (average 6 sec) and ended at the end of a musical phrase. All stimuli were generated from midi files using the Anvil Studio software, with grand piano as instrumental timbre. Videos with the songs' lyrics were created using the “Karafun” software package (Softonic, San Francisco, CA). The lyrics were shown in two or three lines in white Arial font on a black screen. Words changed color from white to red in sync with the rhythm of the corresponding song, in a Karaoke-like fashion that had been judged easy to follow but not too distracting by pilot participants.

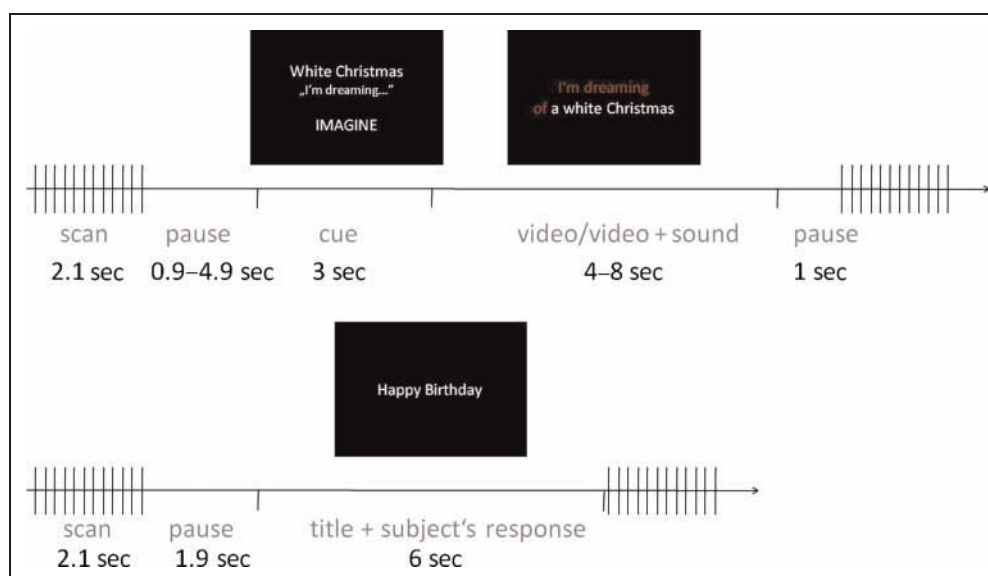
Procedures

Encoding Phase

During the first phase of the experiment, participants listened to or imagined short samples of familiar tunes. On each trial, a screen with the title of the song, the first two words of the lyrics of the specific part used and a cue indicating if the task was to listen to or imagine the song were presented. Thirty Listen and 30 Imagine trials were presented. In Listen trials, participants listened to the song and simultaneously read the lyrics of the song, presented in rhythm with the tune, karaoke fashion. In Imagine trials, the lyrics video was presented without sound, and participants were instructed to imagine the song in time along with the lyrics. In addition to imagine and listen trials, 20 Baseline trials were interspersed during which participants saw a neutral video of 6-sec duration (which was the average length of the melodies used) with a series of “xxxx” instead of words. These word substitutes changed color in an isochronous rhythm, in a karaoke-like style similar to the Listen and Imagery conditions. However, the isochronous rhythm was not expected to cue auditory imagery of any familiar melodies. Baseline trials were cued with the word “Pause,” and participants were instructed to rest their mind but keep their eyes open during these trials. Participants received the instructions for this before the start of the experiment outside the scanner, including some practice trials with tunes that were not used in the real experiment.

Presentation of the stimuli was timed on each trial so that it would end just before the scan started (Figure 1). During scanning and before the presentation of the next cue, participants gazed at a fixation cross in the middle of the screen. Trials were presented in a pseudorandomized order that was different for each subject. Participants were instructed to press a button on a MRI-compatible button box to indicate if they did not know a tune at all. Seven of ten participants indicated this for up to five songs. For each individual subject, unknown songs were excluded from the analysis. Auditory stimuli were presented using

Figure 1. Example trials of the encoding phase (top) and recognition phase (bottom).



scanner-compatible earplugs with foam inserts, and the volume was set to a loudness that was comfortable for the participant (approximately 60 dB) before the start of the scanning.

Recognition Test

After a pause of approximately 30 min during which anatomical scans were acquired, an unannounced recognition test followed. On each trial, participants were shown the title of a song, and they had to indicate via button press on a scanner-compatible button box if they thought that they had listened to this song earlier, or imagined this song earlier, or that this song was new. Songs were not played at all during the recognition phase. As the recognition test was unannounced, participants received the instructions for this part of the experiment including three practice trials directly before the scanning while lying in the scanner.

The assignment of buttons to responses was the same for all participants, and it was shown in the bottom of each slide. Thirty trials were presented in each of the three conditions. Participants had up to 6 sec to give their response, to give them enough time for the dual old/new and source memory decision and to optimize the timing of the scan with respect to the memory decision within our sparse sampling design. They were not specifically instructed to use imagery for solving the task. During 20 Baseline trials, the word "Pause" was shown, and participants were instructed to press any button to control for motor responses. The order of trials was randomized for each subject.

Scanning Protocol

We used a sparse sampling protocol (Belin, Zatorre, Hoge, Evans, & Pike, 1999) both for the encoding and for the

recognition phase. This ensured that the noise of the scanner itself did not interfere with the presentation of the auditory stimuli or with the auditory imagery and prevented artifacts related to responses to the scanner noise itself. EPI images covering the whole head (voxel size = 3.5 mm³, 40 slices, echo time = 30 msec, acquisition time = 2.1 sec, repetition time for the encoding phase = 14 sec, and repetition time for the recognition test phase = 10 sec) were acquired on a Siemens 3T Scanner using a 32-channel head coil. For registration of the functional images to Montreal Neurological Institute (MNI) standard space a high-resolution T1-weighted image with 1 mm³ voxel size was acquired for each participant.

Data Analysis

Behavioral Data

For each participant, errors in old–new discrimination were computed. Then source memory errors were assessed by the number of errors mistaking a previously heard tune for imagined on the recognition test (heard–imagined errors, HI) and the number of errors mistaking a previously imagined tune as heard (imagined–heard errors, IH). Correlations were computed to assess the relation of HI and IH errors, and the relation of auditory imagery vividness and musical expertise with either kind of memory errors.

Functional Data

Functional data were analyzed using the FMRIB software library package (FSL, Oxford, UK). For each individual, functional data of the encoding and recognition phase were registered to the individual high-resolution anatomical image using linear registration and then registered to the MNI 152 standard brain for the group analysis using nonlinear registration. Functional data were

motion-corrected using the motion correction in FSL and spatially smoothed using a kernel of 5 mm FWHM.

Basic analysis of functional data. For the comparison of activity related to the different conditions on the first level, we set up a general linear model including the two different tasks (Listen, Imagine) as explanatory variables in the model for the encoding phase, and including the three different types of stimuli (Listened, Imagined, New) as explanatory variables in the model for the recognition test phase. On the group level, we used a general linear model with explanatory variables group, vividness of imagery (BAIS score part one, vividness of auditory images), and musical training (number of years of formal musical training). For the encoding phase, we looked at the group effects for the contrasts [Listen > Baseline], [Imagine > Baseline], [Imagine > Listen], and [Listen > Imagine] as well as the regression of those contrasts on the vividness scores. For the recognition test phase, analogously, we looked at the contrasts [Previously Listened > Baseline], [Previously Imagined > Baseline], [Previously Imagined > Previously Listened], and [Previously Listened > Previously Imagined] and the regression of those contrasts on the vividness scores. In a further analysis of the encoding and recognition phases, we compared trials that resulted in HI and IH errors with the respective correct trials of these conditions. In all of these analyses, the amount of training was taken into account as a factor in the general linear model (number of years of formal musical training), so that the results reflect activity independent of musical training background. Significant activation was determined using the cluster threshold in FSL. The initial z threshold for clusters was set at $z = 2.3$. The cluster-corrected level of significance was set at $\alpha = 0.05$ for all analyses. Because we used the MNI 152 target for registration, coordinates are reported in MNI space.

Conjunction analysis. Activity that was common to both imagining and listening (encoding phase) and to remembering previously heard and previously imagined tunes (recognition test) on the group level was determined by the intersection of significant activity in the corresponding contrasts to baseline. However, finding overlapping activity on a group level does not necessarily imply overlapping activity in individuals as well. Therefore, we additionally did the conjunction analysis on an individual level and then determined how many participants showed overlap in each voxel after transformation to the standard space.

Functional connectivity analysis. To determine if functional connectivity during imagery and perception differed, we conducted analyses of functional connectivity and its modulation by task demands, sometimes referred to as psychological–physiological interaction (Friston et al., 1997). Seed voxels for the connectivity analysis were based on the results of the conjunction. We selected the

local maxima in anterior and posterior auditory areas in the superior temporal gyrus (STG) of both hemispheres as well as in the SMA, in the clusters that showed activity both during imagery and listening. These seed voxels were projected back from standard space to individual space, and the time series of BOLD response in these voxels was entered as an explanatory variable in the general linear model on the first level for each individual. The functional connectivity of activity in this area, depending on the task, was determined by the interaction of the BOLD time series with a vector that incorporated the contrast of the two conditions ($A > B$). To account for the shared variance, an additional explanatory variable that includes both conditions ($A + B$) was included in the model. The group level analysis was conducted analogously to the basic analysis of the functional data.

RESULTS

Behavioral Data

Overall old–new recognition performance in the recognition task was good (mean probability of old–new recognition errors = 0.17, $SD = 0.06$). The mean probability of HI errors (mistaking a previously heard song for an imagined song) was 0.23 ($SD = 0.17$); mean probability of IH errors was comparable at 0.24 ($SD = 0.13$). However, error rates for the two kinds of source memory failures (HI and IH errors) were not significantly related (Pearson's $r = -.18$).

Our prediction was that strong auditory imagery vividness should lead to more source memory errors. The number of HI errors correlated negatively with auditory imagery vividness scores ($r = -.69$, $p = .027$): The lower the vividness score, the more errors mistaking previously heard melodies for imagined in the recognition task. In contrast, the correlation of the number of IH errors and imagery vividness was low ($r = -.102$, ns). No significant correlations were found for false memory errors and the amount of musical training that participants had received ($r = .07$ for HI errors, and $r = .03$ for IH errors).

Functional Data

Encoding Phase

Activity related to listening and imagining familiar songs during the encoding phase of the experiment was assessed in the contrasts [Listen > Baseline] and [Imagine > Baseline]. Figure 2A shows both contrasts as well as the conjunction of the effects on the group level. Figure 2B shows the conjunction results based on individual conjunctions in each subject.

During listening, not surprisingly, primary and secondary auditory cortices are active. Based on visual inspection none of the areas of conjunction overlapped with Heschl's gyrus and immediately adjacent areas, and therefore, we conclude that primary and adjacent secondary auditory

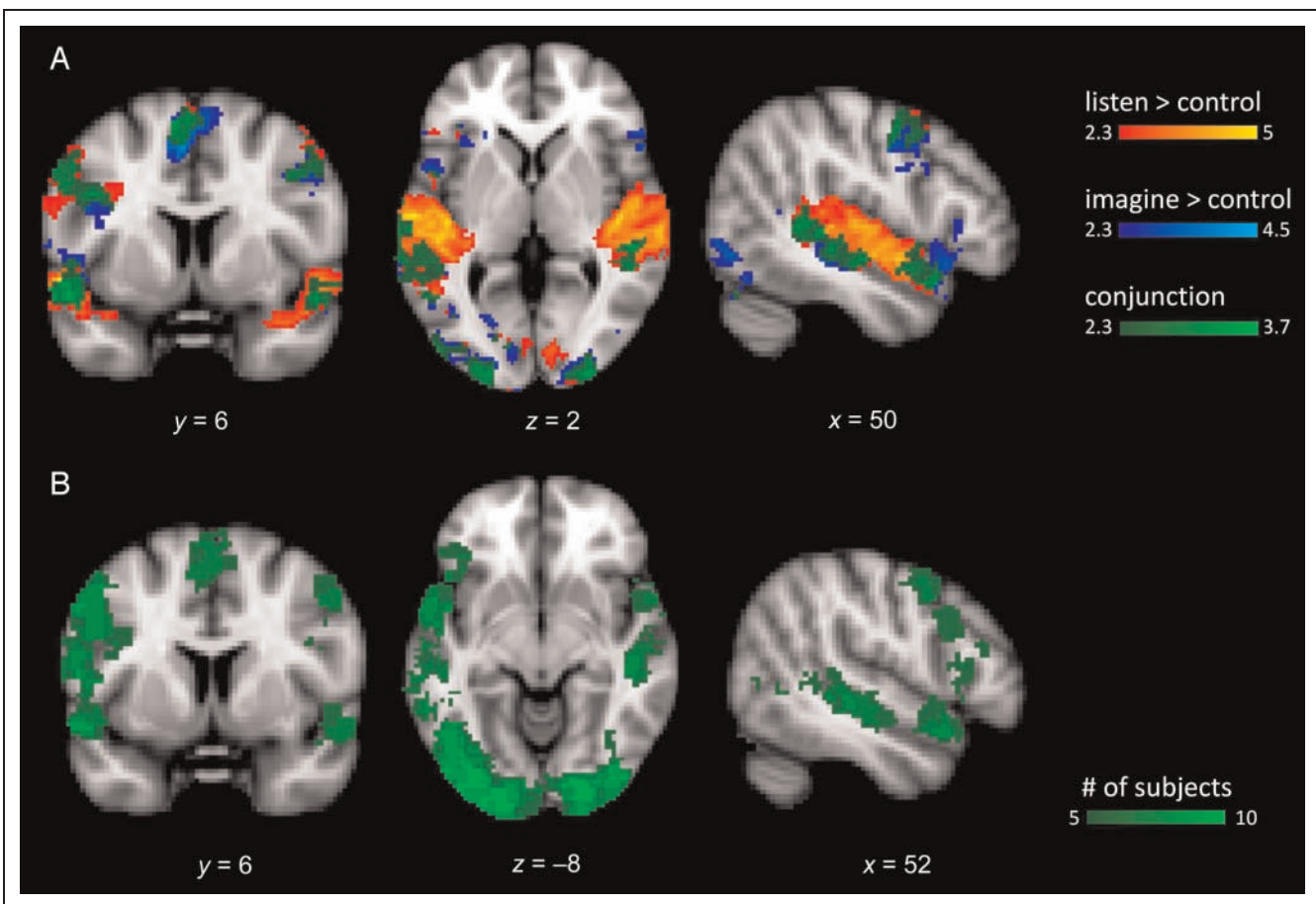


Figure 2. (A) Activity during the encoding phase of the experiment during listening to (red) and imagination (blue) of familiar melodies compared with baseline (cluster threshold in FSL with $z > 2.3$, $p < .05$). Green areas show the overlap of the activation maps. (B) Overlap of activity during imagining and listening compared with baseline was also computed on an individual level. Green areas show voxels where at least five participants showed significant activity in both conditions compared with baseline.

cortices were not significantly active during imagery. However, we found activity in anterior and posterior auditory association areas in the STG in both hemispheres, overlapping with the same areas that are also active during listening, as is evident from the conjunction of the results. Additional areas that are active during both conditions are large clusters covering the premotor cortex bilaterally and extending to inferior frontal gyrus on the left side, SMA, right OFC, and large parts of the visual cortex. Also, we found extended overlapping activity in the cerebellum with a peak in left lobule VI, but extending to crus I and II, lobules VI, VIIb bilaterally, vermis and right VIIIa. Peak voxels of significant activations are given in Table 1.

A direct comparison between the conditions that reveals differences between activity related to imagining and listening is shown in the contrasts (Imagine > Listen) and (Listen > Imagine) in Figure 3. Primary and secondary auditory areas bilaterally, including Heschl's gyrus and immediately surrounding STG and planum temporale, show stronger activity during listening than during imagery. Also, a small cluster in the precuneus region was more active during listening. In contrast, a number of areas are more strongly activated during imagery compared

with listening. These areas include the IPS extending to supramarginal gyrus bilaterally, the SMA extending to paracingulate gyrus and ACC, parts of dorsolateral pFC (DLPFC) bilaterally, precentral gyrus extending to middle frontal gyrus on the right hemisphere, insular cortex bilaterally extending to frontal operculum cortex on the right hemisphere. Also, we found increased activity during imagery compared with perception in the cerebellum in the left lobule VI extending to crus I, which overlaps with activity seen during both conditions in the conjunction. Notably, no areas within the STG that could be considered as auditory were more active during imagery than during listening.

Using participants' BAIS vividness scores as an explanatory variable in a regression analysis on the contrast (Imagery > Baseline) revealed a relationship between the general capacity for vivid auditory imagery and brain activity during imagery. Two areas are especially active during imagery in participants with higher auditory imagery vividness (Figure 4): part of right DLPFC and an anterior part of the right STG. The cluster in the STG does not overlap with the conjunction in the group average activity during both imagery and perception of the tunes (cf. Figure 2),

Table 1. Peak Voxels of Significant Clusters during the Encoding Phase of the Experiment as Determined by the Cluster Threshold in FSL ($z > 2.3$, $p < .05$)

<i>Anatomical Location (According to Harvard Oxford Atlas in FSL)</i>	<i>Voxels</i>	<i>Z-MAX</i>	<i>Z-MAX x (mm)</i>	<i>Z-MAX y (mm)</i>	<i>Z-MAX z (mm)</i>
<i>Listen > Baseline</i>					
Left planum temporale	12212	5.04	−62	−16	4
Right STG/Heschl's gyrus	4011	4.85	46	−16	0
Left precentral gyrus	1590	3.84	−58	0	44
Left intracalcarine cortex	270	3.54	−6	−78	10
SMA	254	3.37	−8	4	58
Right precentral gyrus	236	4.05	54	0	48
<i>Imagine > Baseline</i>					
Left occipital fusiform gyrus	10488	4.53	−46	−64	−18
Left precentral gyrus	2600	3.77	−56	2	42
Superior frontal gyrus/SMA	852	3.99	−2	10	58
Right STG	710	3.89	54	−14	−10
Lateral occipital cortex	537	3.92	−36	−60	58
Right precentral gyrus	490	3.85	54	0	48
Right temporal pole	446	3.56	56	16	−18
<i>Conjunction of [Listen > Baseline] and [Imagine > Baseline]</i>					
Lingual gyrus/occipital fusiform gyrus	7379	4.32	−16	−76	−12
Precentral gyrus	1284	3.74	−56	2	42
Right STG	597	3.89	54	−14	−10
Left STG	489	3.76	−52	8	−12
SMA	245	3.37	−8	4	58
Right precentral gyrus	204	3.85	54	0	48
Right STG	204	3.19	52	6	−14
Intracalcarine cortex	133	3.27	−8	−80	10
<i>Regression with Vividness on Contrast Imagine > Baseline</i>					
Right frontal pole	414	3.49	18	50	26
Right anterior STG	274	3.54	52	4	−4
<i>Listen > Imagine</i>					
Left Heschl's gyrus	4335	5.13	−40	−26	4
Right inferior frontal gyrus	4227	5.43	52	−12	2
Precuneus cortex	266	2.95	2	−62	20
<i>Imagine > Listen</i>					
Paracingulate gyrus (extending to SMA and anterior cingulate cortex)	3354	3.62	2	12	48
Right insular cortex	2242	3.76	46	14	−4

Table 1. (continued)

Anatomical Location (According to Harvard Oxford Atlas in FSL)	Voxels	Z-MAX	Z-MAX <i>x</i> (mm)	Z-MAX <i>y</i> (mm)	Z-MAX <i>z</i> (mm)
Right superior parietal lobule	1164	3.1	34	−50	50
Angular gyrus	691	3.03	−56	−54	30
Left frontal pole	615	3.26	−32	44	28
Cerebellum: left VI/left Crus I	407	3.41	−34	−54	−32
Left insular cortex	374	3.56	−36	16	−2

For each cluster, the size in number of voxels, the peak *z* value (Z-MAX), and its coordinates in MNI space in millimeters (Z-MAX *x*, *y*, and *z*) are given. Conjunction of contrasts was computed as the intersection of significant areas (minimum of both *z* values). Estimated anatomical locations were determined from the Harvard Oxford Atlas implemented in FSL.

but is within an area that is overall only significantly active during listening. The cluster in right pFC partly overlaps with the areas that were found to be overall increased during imagery compared with perception but does not overlap with activity that is found during imagery compared with baseline in all participants.

Functional Connectivity

In the functional connectivity analysis, the only seed area among the four auditory areas and the SMA that showed a significant difference in connectivity was the seed in right anterior STG ($x = 52, y = 6, z = -14$). From this area, connectivity to right prefrontal areas (Figure 5) was increased during imagery compared with during perception. To reveal the source of this interaction, we computed the functional connectivity for both of the basic contrasts ([Imagine > Baseline] and [Listen > Baseline]) and confirmed that this fronto-temporal connectivity was only present in the contrast of the imagery condition versus baseline. The areas that are more strongly connected during imagery compared with perception in the interaction overlap with the activity specifically related to imagery compared with perception in the initial analysis of the

functional data. No areas were significantly more correlated with activity in the seed voxels during perception compared with imagery. If the statistical threshold is lowered, similar patterns of enhanced connectivity during imagery compared with perception from the STG regions on both hemispheres, and from the SMA to prefrontal areas emerge, indicating that this is a consistent pattern of results across these regions.

Recognition Phase

During recognition, the pattern of BOLD activity was similar during recognition of heard and imagined songs (Figure 6A) in inferior frontal cortex and ACC, SMA, pFC, as well as in the left middle temporal gyrus. The direct contrast of both conditions did not reveal any difference between recognition of previously imagined and previously heard melodies.

Whereas regression with BAIS vividness scores during recognition of imagined songs did not reveal any significantly activated areas, we did find activity significantly related to vividness of imagery during recognition of previously heard songs. This area is part of the left frontal pole (Figure 6B) but does not overlap with areas that were

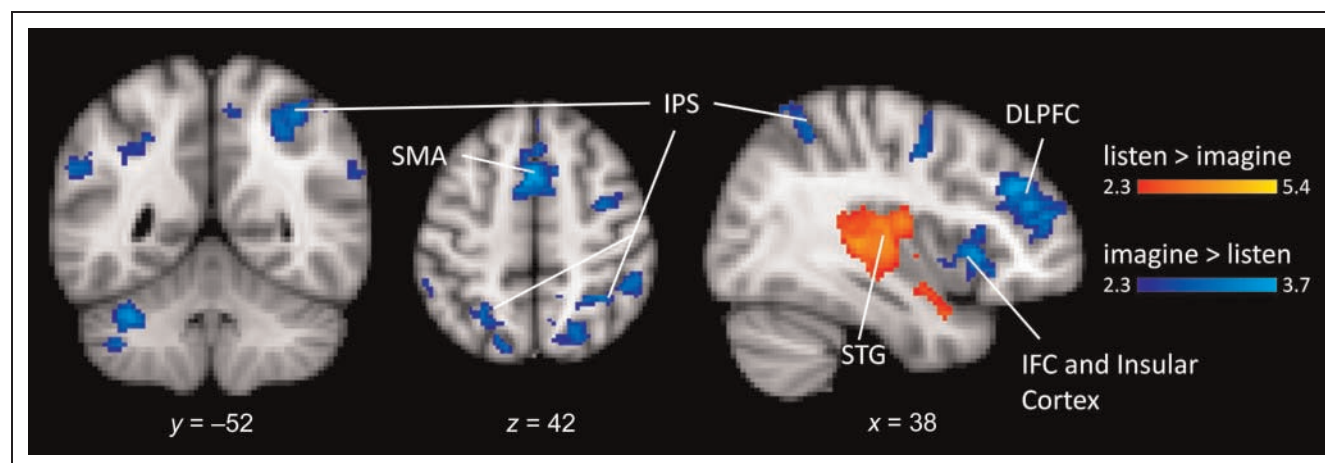


Figure 3. Direct comparison of activity during listening to (red) and imagination (blue) of familiar melodies during the encoding phase of the experiment (cluster-threshold in FSL with $z > 2.3, p < .05$).

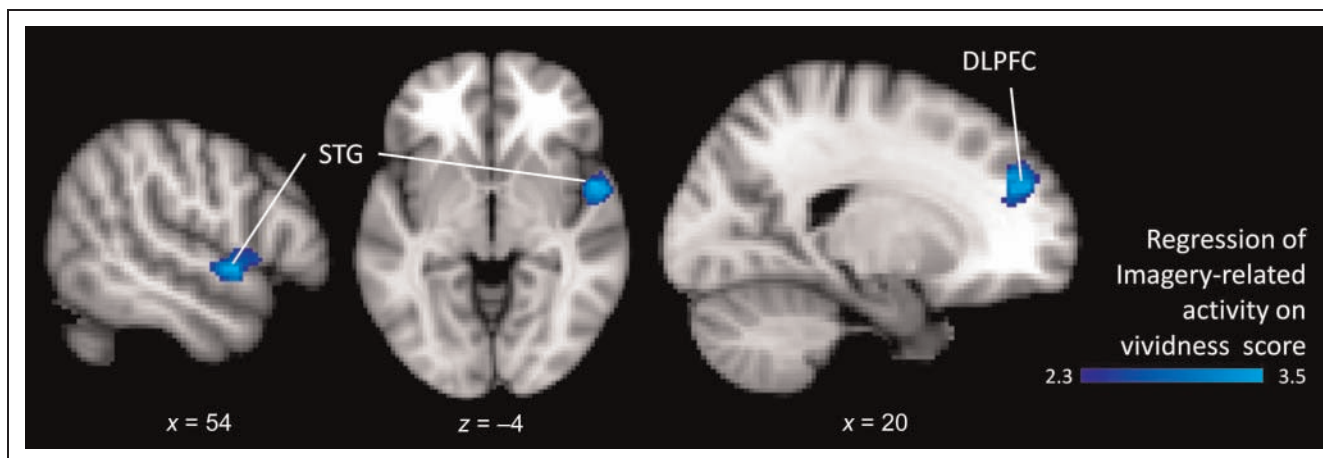


Figure 4. Regression of activity during imagery compared with baseline during the encoding phase of the experiment on the vividness of imagery as assessed by the BAIS score (cluster threshold in FSL with $z > 2.3$, $p < .05$).

more active during imagery than during perception during the encoding phase of the experiment. Table 2 shows the locations of the peak voxels for the recognition phase.

A further analysis of the functional data based on the participants' responses in the recognition task, comparing activity for trials during encoding that are later either correctly or incorrectly remembered did not yield any significantly activated areas. This might be because of the low number of trials that were available for this comparison for most participants, which results in low statistical power to detect effects.

DISCUSSION

The Role of Auditory Cortices for Auditory Imagery during Encoding

Consistent with previous studies on imagery and perception of music, we have shown a large extent of overlap in

activity in auditory association areas during those two processes (Zatorre & Halpern, 2005; Halpern et al., 2004; Schürmann et al., 2002; Halpern & Zatorre, 1999; Zatorre et al., 1996). In comparison with previous studies, we have shown both overlap in anterior and posterior STG. Also, we were able to show that this overlap is not only present on the group level (Figure 2A) but also if the conjunction is done on an individual level (Figure 2B). This detail is important in showing that the overlap is not an artifact of averaging. Although this type of overlap has been postulated in previous studies on musical imagery, the high resolution of the present data and the conjunction results show a clearer and more detailed picture of how auditory areas are involved not only in perception of music but also in the internal generation of auditory images. Our results show a distribution of activity throughout anterior and posterior STG that represents a subset of areas responsive to sound. Although some studies have reported involvement of primary auditory cortex (Kraemer, Macrae, Green,

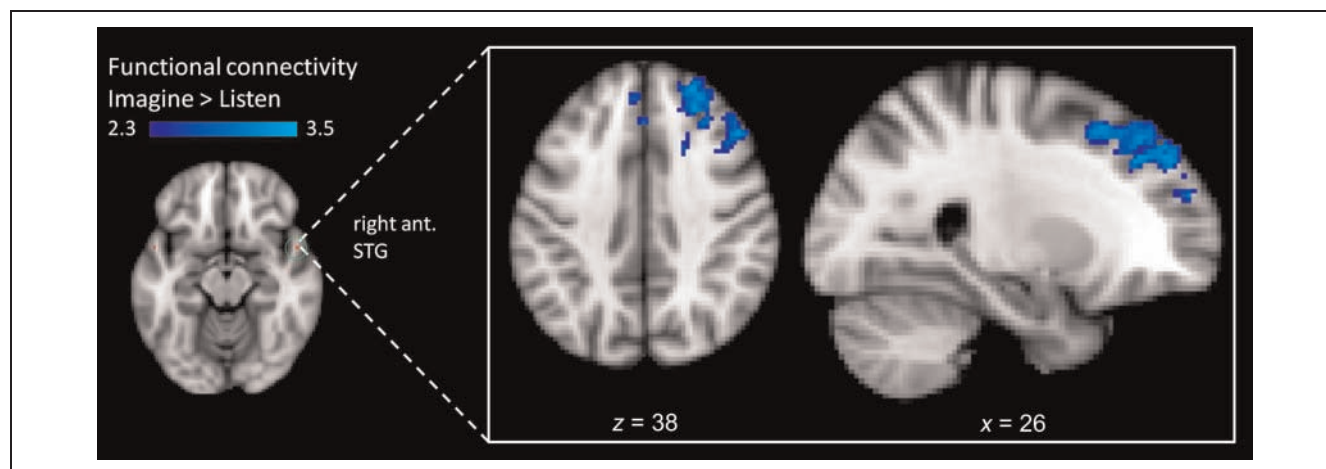


Figure 5. Results of the analysis computing changes in functional connectivity (PPI), showing the increase in connectivity from the seed in the anterior part of the right STG (small inset) to right prefrontal areas in Imagine as compared with Listen conditions (cluster threshold in FSL with $z > 2.3$, $p < .05$).

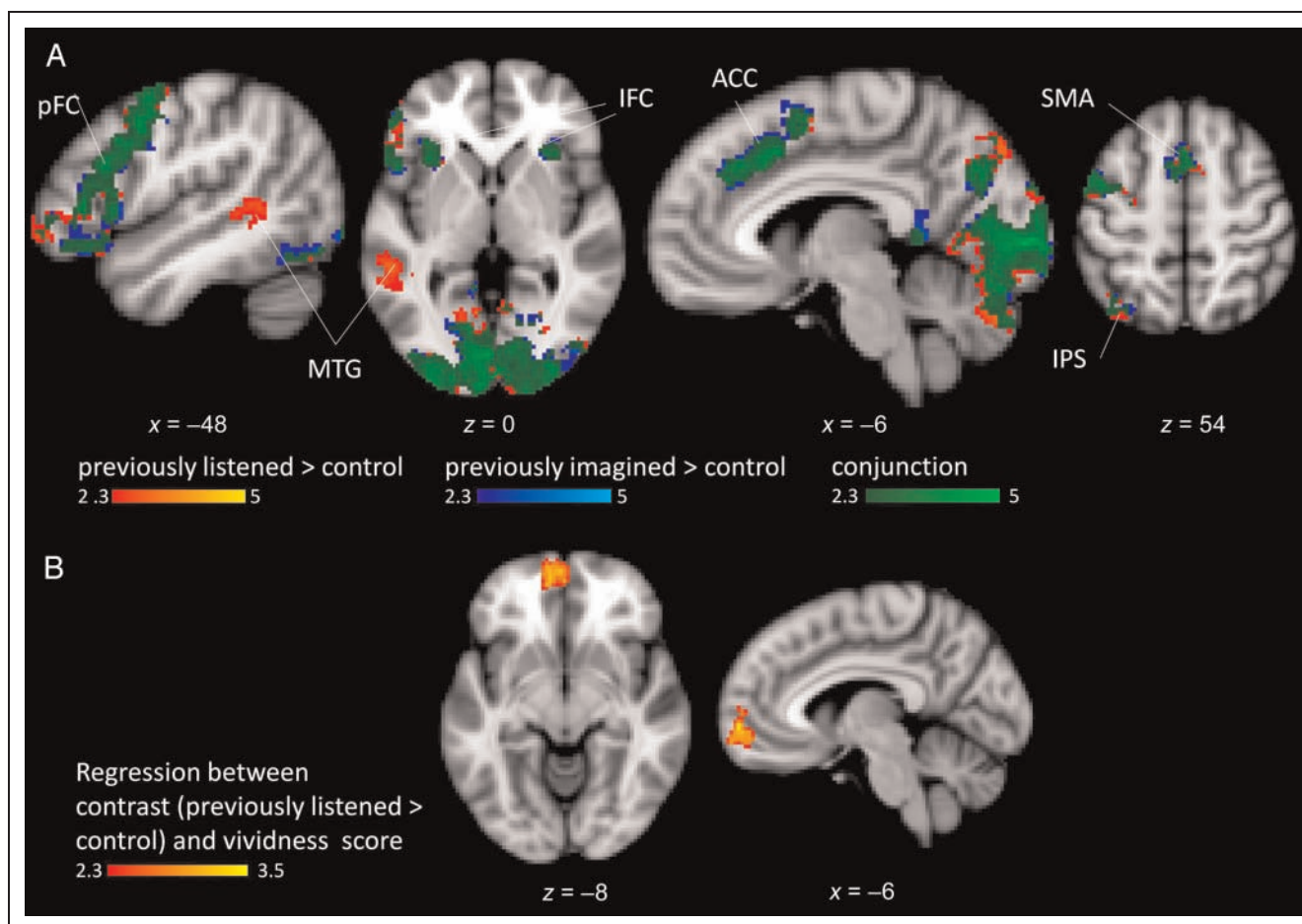


Figure 6. (A) Activity during the recognition test phase of the experiment during recognition of previously listened (red) and previously imagined (blue) familiar melodies compared with baseline (cluster threshold in FSL with $z > 2.3$, $p < .05$). Green areas show the overlap of the activation maps. (B) A regression of the activity during recognition of previously heard songs compared with baseline on BAIS vividness scores revealed significant activity in the left frontal pole.

& Kelley, 2005; Yoo, Lee, & Choi, 2001), this is still a topic under debate, as these studies did not provide independent identification of primary areas. Consistent with our previous studies (Zatorre et al., 1996, 2010; Halpern et al., 2004; Halpern & Zatorre, 1999), we failed to find primary auditory cortex involvement above statistical threshold.

The correlation of a person's general ability to vividly imagine auditory information (BAIS score) with brain activity during imagery compared with baseline revealed a cluster in the anterior part of the right STG. More vivid imagers showed significantly more activity in this area while performing our auditory imagery task than people who reported only weak imagery capabilities. Interestingly, this cluster did not overlap with activity in the group average during imagery but does overlap with activity during listening in the whole group. This seems to indicate that more vivid imagers tend to recruit an auditory area that is involved in perception of tunes more strongly during the imagery of these tunes. Previous studies that have looked at correlations with vividness have correlated brain activity with trial-by-trial ratings of subjective vividness. Whereas

Leaver et al. (2009) did not find correlations with auditory areas in a whole-brain search, Daselaar et al. (2010) showed a relation of brain activity during imagery in auditory areas with on-line vividness ratings. However, because of the statistical approach, their conclusions of vividness-related increase of activity were restricted to areas that were generally active during very vivid imagery. We show that even scores from an off-line questionnaire that is not specific to the stimuli used in the experiment predict activity in auditory areas during imagery. This finding is consistent with research from vision, where increased activity in visual areas during visual imagery has been found in vivid imagers (Cui et al., 2007), and suggests that an enhancement of activity in modality-specific areas as a characteristic of people with more vivid imagery might be a domain-general mechanism.

Importantly, all our conclusions about recruitment of auditory areas take musical experience of participants into account. As it has previously been shown that musical experience modulates brain correlates of auditory (Herholz et al., 2008) and auditory-motor imagery (Lotze, Scheler, Tan, Braun, & Birbaumer, 2003), we made sure to have a

sample of participants whose auditory vividness scores were only very weakly related to musical training background, and we used years of musical training in our analyses as a covariate. Therefore, our results can be considered independent of musical expertise, showing that not only this factor (Herholz et al., 2008; Lotze et al.,

2003) but also the general ability of vivid auditory imagery influences brain correlates of auditory imagery. The lateralization of the activity to the right hemisphere during more vivid auditory imagery of familiar tunes in this study is especially interesting as the melodies included lyrics. We assumed that left-lateralized language processing

Table 2. Peak Voxels of Significant Clusters during the Recognition Phase of the Experiment as Determined by the Cluster Threshold in FSL ($z > 2.3$, $p < .05$)

<i>Anatomical Location (According to Harvard Oxford Atlas in FSL)</i>	<i>Voxels</i>	<i>Z-MAX</i>	<i>Z-MAX x (mm)</i>	<i>Z-MAX y (mm)</i>	<i>Z-MAX z (mm)</i>
<i>Previously Listened > Baseline</i>					
Occipital pole	10382	4.8	−6	−92	2
Left insular cortex	3235	4.28	−32	20	−6
Paracingulate gyrus	1088	4.13	−4	32	36
Right inferior frontal gyrus	381	4.12	46	20	20
Right frontal orbital cortex	353	4.31	30	20	−10
Cerebellum: Right VIIb	302	3.62	34	−70	−50
Left STG	255	3.49	−56	−34	2
Left angular gyrus	253	3.41	−38	−58	44
<i>Previously Imagined > Baseline</i>					
Occipital pole	11298	4.97	−6	−96	4
Precentral gyrus	3409	4.73	−40	8	28
Paracingulate gyrus	1260	4.14	−8	20	40
Right insular cortex	424	4.47	32	22	−4
Right inferior frontal gyrus	279	3.82	46	20	20
<i>Conjunction of [Previously Listened > Baseline] and [Previously Imagined > Baseline]</i>					
Occipital Pole	8978	4.8	−6	−92	2
Left inferior frontal gyrus	2881	4.27	−38	10	26
Paracingulate gyrus	1022	3.93	−8	20	40
Left frontal orbital cortex	347	4.14	30	20	−10
Right inferior frontal gyrus, pars opercularis	253	3.82	46	20	20
Cerebellum: right VIIb	236	3.56	32	−66	−52
Right angular gyrus/sup, par, lobule	188	3.23	−38	−56	44
Posterior cingulate gyrus	60	3.11	−2	−50	10
Lingual Gyrus	13	2.5	8	−60	0
<i>Regression with Vividness on Contrast Previously Listened > Baseline</i>					
Left frontal pole	328	3.35	−6	60	−2

For each cluster, the size in number of voxels, the peak z value (Z-MAX), and its coordinates in MNI space in millimeters (Z-MAX x , y , and z) are given. Conjunction between contrast was computed as the intersection of significant areas (minimum of both z values). Estimated anatomical locations were determined from the Harvard Oxford Atlas implemented in FSL.

was occurring in this condition. Indeed, in the overall activity (group mean) during imagery (Figure 2) more widespread activity was observed in left auditory areas, covering Wernicke's area on the left hemisphere. However, the vividness of the auditory images seems to be more strongly related to a right-hemispheric network. This is in line with findings of right-lateralized processing of pitch information and melodies (Zatorre, Belin, & Penhune, 2002) and suggests that it is the tonal aspect of melodic imagery that is reflected in this result.

Other Areas Common to Both Imagery and Perception

In our experimental setup, the presentation of the lyrics in a karaoke-like fashion strongly encouraged mental singing. Consistent with this task characteristic, premotor cortices bilaterally were active during imagery as well as perception, but were not enhanced in either condition compared with the other, indicating that these areas do not specifically relate to either activity. Co-activation of motor and premotor areas during listening and during imagery has also been shown in other studies (Kleber, Birbaumer, Veit, Trevorrow, & Lotze, 2007; Lahav, Saltzman, & Schlaug, 2007; D'Ausilio, Altenmüller, Olivetti Belardinelli, & Lotze, 2006; Lotze et al., 2003) and the location of our finding is consistent with the location of the representation of the larynx in other studies (Brown, Ngan, & Liotti, 2008; Loucks, Poletto, Simonyan, Reynolds, & Ludlow, 2007), suggesting that part of the recruited motor network may be related to inner singing.

Our data further support the important role of the auditory-motor loop in musical imagery and perception (Zatorre, Chen, & Penhune, 2007), even when controlling for musical experience. In the left hemisphere, activity common to imagery and perception extended to Broca's area in the inferior frontal gyrus, in line with speech processing taking place both during listening and imagining (Kleber et al., 2007). Mental singing of the tunes and reading of the lyrics in both conditions provide a plausible explanation for the involvement of language-related areas. Similarly, we found strong and extensive activity in visual cortices during imagery and perception compared with baseline. This is probably reflecting stronger visual processing during the reading of real tune lyrics compared with our baseline stimulus. However, in the direct contrast, perception and imagery did not differ significantly in the visual areas, indicating that the visual processing was similar in both conditions.

Functional Cortical Network Underlying Auditory Imagery

We directly contrasted activity during imagery and perception with otherwise comparable task conditions, showing that whereas perception activates bilateral primary and secondary auditory areas more strongly, activity during

imagery is stronger in a network of frontal and parietal areas and cerebellum. Whereas activity in other parts of the motor network was similar during imagery and perception, we found increased activity in the cerebellum (Lobule VI) during imagery, an area that contains the representations of tongue and lip movements (Grodde, Hülsmann, Lotze, Wildgruber, & Erb, 2001). Our findings are consistent with findings of bilateral Lobule VI activity during mental singing in professional singers (Kleber et al., 2007). Differences in extent and lateralization between studies might be related to group characteristics, different task demands of explicit instruction to imagine the motor act of singing in Kleber et al. (2007), versus more general instructions to imagine the song, and statistical thresholding, as a homologous cluster in the right cerebellar hemisphere was just below threshold in our data. Our results further support the cerebellar contribution to subvocalization during imagery, even in participants with more limited amounts of musical experience.

The SMA was involved both during listening and during imagining of the melodies. However, it was involved to a larger extent during the imagery task, in line with previous studies on auditory imagery (Halpern & Zatorre, 1999; Zatorre et al., 1996). Although in the context of this experiment the SMA is probably part of the network supporting mental singing and subvocalization, as suggested by its subthreshold appearance in the functional connectivity analysis, it has also been shown to be active during the imagery of sounds that are difficult to vocalize with the human voice, such as instrumental timbres (Halpern et al., 2004), indicating a role for the SMA in mental imagery over and above the obvious vocal motor component.

In our results, the IPS was more active during imagery than during perception bilaterally. Our stimulation differed from previous studies in that we used a karaoke-like presentation of the lyrics that participants followed both during imagery and listening. Although the baseline condition did not require any actual reading, the increased IPS activity was found between the imagery and listen conditions, which were comparable regarding the presented stimuli. Interestingly, bilateral parietal regions including the IPS have been found to support simultaneous attention to spatial locations and temporal intervals (Coull & Nobre, 1998), cross-modal detection of motion in the visual and auditory domain simultaneously (Lewis, Beauchamp, & DeYoe, 2000), and audio-visual integration of non-speech stimuli (Calvert, Hansen, Iversen, & Brammer, 2001), pointing to the role of the IPS in cross-modal integration and attention. In the context of our experiment, matching the visual input of the karaoke-style lyrics with the imagined tunes might have demanded more resources than merely perceiving simultaneous lyrics and tunes during the listening condition. Our finding of increased IPS activity during auditory imagery while synchronizing with a dynamic visual input could reflect audio-visual cross-modal integration, not only during perception but also during mental imagery.

In our analyses, the pFC featured prominently. We found an overlap of regions or close spatial proximity in the activations specific to imagery (compared with perception) and to vividness of imagery in right DLPFC, pointing to its role in the imagery of familiar music. Right DLPFC was specifically more active during auditory imagery in those participants that had a generally higher capacity for vivid auditory imagery. Daselaar et al. (2010) showed that activity in pFC was correlated with imagery vividness ratings on a trial-by-trial basis both for auditory and visual imagery. In combination, these findings indicate that activity in DLPFC is related to imagery vividness both as a state and as a trait. Interestingly, the cluster in our analysis partly, but not completely overlapped with areas that are specific to auditory imagery across the whole group, indicating that not only right auditory areas but also right DLPFC are additionally recruited in more vivid imagers, and suggesting that their superior imagery capabilities rely on activity in this right-lateralized network.

Fronto-temporal Connectivity in Mental Imagery

The important role of the pFC in auditory imagery is also evident from the results of the functional connectivity analysis, where we found increased correlation of right anterior temporal STG with predominantly right prefrontal cortical areas. Halpern and Zatorre (1999) showed that imagery of real, familiar tunes was related to activation primarily in right frontal areas and right STG. Here we show that these areas are not only important for imagery, but that they are moreover part of a functional network, as evident in the results of the functional connectivity analysis. In visual imagery, increased connectivity from prefrontal to visual areas compared with perception has been demonstrated as well. Although functional connections between visual areas and parietal cortex are not modulated by stimulus content, the prefrontal top-down connection to visual areas apparently conveys content-related information during imagery (Mechelli et al., 2004). In combination with our results, this indicates that the pFC has a domain-general role in the interaction with the specific sensory areas involved in the imagery modality. Possible roles of the frontal cortex during auditory imagery are memory retrieval, working memory and mental monitoring (Petrides, 2000; Henson, Shallice, & Dolan, 1999). In one of our previous studies on musical imagery, we have shown that frontal areas are specifically involved in the retrieval of the familiar melodies from long-term memory (Halpern & Zatorre, 1999). However, the areas in this previous study were more inferiorly located than the present results. pFC has also been implicated in tonal working memory (Koelsch et al., 2009; Zatorre, Evans, & Meyer, 1994), and psychological research has provided evidence that the phonological or tonal loop is involved in auditory imagery both for familiar and unfamiliar material (Baddeley & Andrade, 2000). One important feature of voluntary auditory imagery is that in contrast to auditory hallucinations,

the person is aware that the auditory impression is self-generated and controllable, which requires monitoring of the self-generated auditory impression. Functional connectivity of frontal and auditory areas has previously been shown to be diminished in schizophrenic patients suffering from auditory hallucinations (Lawrie et al., 2002) and seems to play a role also in the experience of auditory hallucinations in patients with acquired deafness (Griffiths, 2000). Such clinical observations combined with our findings point to a role of the functional temporo-frontal connection for the mental monitoring of internally generated auditory impressions that seems to be impaired during auditory hallucination, and that is enhanced during active, voluntary auditory imagery in healthy individuals.

Recognition Test

In the second part of this study, we presented an unannounced recognition test to the participants, in which they had to judge based on the titles of the songs if during the first part of the experiment they had heard the song, imagined it, or not heard it at all. On the behavioral level, we were interested in source memory errors where participants would mistakenly remember a heard song as imagined or vice versa, and how these errors were related to the participants' general vividness of imagery. We found a significant correlation with vividness for errors where participants mistook a heard for an imagined song, with participants who had very low vividness scores making significantly more such errors than those who were more vivid imagers. In contrast, in the visual domain, more vivid imagery has been found to impair performance in source memory tasks (Eberman & McKelvie, 2002; Dobson & Markham, 1993). However, as those studies compared memory for items that were both physically presented, albeit in different modalities (either as film/audio or text), a direct comparison with our results is difficult. A possible interpretation of our findings in the context of this study is that lower imagers encode and store with weaker perceptual-like qualities, and therefore more tunes would seem to be imagined to them at retrieval.

A second question related to the recognition phase was which neuronal correlates support remembering heard and imagined events. Previous studies showed that not only actively imagining but also remembering auditory information can result in activity in secondary auditory areas (Huijbers et al., 2011; Wheeler et al., 2000). However, in contrast to our result for the imagery-related activity during the encoding phase, in the recognition phase we did not find significant activity in auditory areas during the recognition of previously heard or imagined tunes. We did find activity in the left STS during the recognition test. Although this area turned out significant only for previously heard stimuli, lowering the threshold showed a similar area also for previously imagined tunes. This activity during the recognition test could be a correlate of auditory imagery during the memory retrieval, but might possibly

also be related to language processing due to reading or remembering the titles of the songs. Memory for text and melody of songs are intimately related (Wallace, 1994; Serafine, Davidson, Crowder, & Repp, 1986), and in the context of this experiment, it is difficult to make assumptions as to which aspect of the stimuli was predominantly remembered. However, the left middle temporal gyrus has also been shown to be involved in semantic memory of music (Platel, Baron, Desgranges, Bernard, & Eustache, 2003), suggesting that participants might have accessed semantic memory contents for tunes that they had really heard in the preceding phase of the experiment. We chose to use titles as cues for the recognition test, as we were interested in the recall of the source of the memory (imagined or heard). Titles serve as a powerful cue for the recall of melodies, especially for songs with lyrics (Peynircioglu, Tekcan, Wagner, Baxter, & Shaffer, 1998). We cannot exclude the possibility that participants might have solved the task without intense imagery or recall of the melodies by relying only on the titles of the tunes that were also presented during the encoding phase. This might be a possible reason for the lack of other auditory cortex activity during the recollection of auditory information in the context of our experiment. However, our finding of a relationship of imagery vividness with source memory errors seems to suggest that participants attempted to retrieve actual information about the perceptual qualities during the encoding phase and thus at least partly engaged in reimagining of the songs.

Activity in other areas that were found active during recognition of both imagined and heard songs can be related to (episodic) memory retrieval of music for the inferior frontal cortex (Watanabe, Yagishita, & Kikyo, 2008; Platel et al., 2003) and mental imagery processes to help in the decision for pFC and SMA (Zatorre & Halpern, 2005), but these areas might also be engaged due to motor preparation of the response (Richter, Andersen, Georgopoulos, & Kim, 1997; Deiber, Ibañez, Sadato, & Hallett, 1996) and other memory-related aspects of the task (pFC), for example, postretrieval monitoring processes (Schacter, Buckner, Koutstaal, Dale, & Rosen, 1997).

We found part of the left medial pFC especially active during recognition of previously heard tunes in participants who had more vivid auditory imagery. This result is consistent with the behavioral results: More active imagers apparently reinstate the memory of the heard melodies better than low imagers. Medial pFC is implicated in a number of different systems, including emotional processing (Phan et al., 2003), social cognition (Amodio & Frith, 2006), self-referential mental activity (Gusnard, Akbudak, Shulman, & Raichle, 2001), and memory for self (Macrae, Moran, Heatherton, Banfield, & Kelley, 2004), and belongs to a network that is relevant both for prospective imagery and episodic memory of past events (Addis, Pan, Vu, Laiser, & Schacter, 2009). In the context of this study, we interpret our finding as an indication that the overall ability to vividly imagine not only affects neuronal

correlates of active, voluntary auditory imagery, but also supports remembering previous auditory events. A possible reason for the absence of such a relation for previously imagined trials might be availability of a recent auditory memory for heard trials that evoked a stronger auditory image during recall, especially in participants with more vivid auditory imagery capabilities.

Summary and Conclusion

The very clear results regarding the recruitment of secondary auditory areas during auditory imagery that overlap with activity during perception, as well as the cortical network underlying auditory imagery, but not perception of familiar tunes, confirm and extend the findings of previous studies. The new findings of temporal and frontal areas specifically related to vividness of imagery, as well as the finding of increased functional connectivity of right-hemispheric prefrontal and temporal areas during imagery, create a link between the reported subjective experience and brain function, hence providing new insights in the cortical networks underlying mental imagery.

Acknowledgments

We would like to thank David Lizotte for his work on developing the BAIS and Amanda Child, Daniele Gold, Rachel Paston, and Alexandre Apfel for pilot testing the memory task. This research was supported by Bucknell University Scadden Faculty Scholar Award, Deutsche Forschungsgemeinschaft (HE6067-1/1) and CIHR (MOP-14995/MOP-11541) and NSERC.

Reprint requests should be sent to Sibylle C. Herholz, Montreal Neurological Institute, McGill University, Room #276, 3801 rue University, Montreal, Quebec, Canada, H3A 2B4, or via e-mail: sibylle.herholz@mail.mcgill.ca or sibylle.herholz@googlemail.com.

REFERENCES

- Addis, D. R., Pan, L., Vu, M.-A., Laiser, N., & Schacter, D. L. (2009). Constructive episodic simulation of the future and the past: Distinct subsystems of a core brain network mediate imagining and remembering. *Neuropsychologia*, 47, 2222–2238.
- Amedi, A., Malach, R., & Pascual-Leone, A. (2005). Negative BOLD differentiates visual imagery and perception. *Neuron*, 48, 859–872.
- Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, 7, 268–277.
- Baddeley, A., & Andrade, J. (2000). Working memory and the vividness of imagery. *Journal of Experimental Psychology: General*, 129, 126–145.
- Belin, P., Zatorre, R. J., Hoge, R., Evans, A. C., & Pike, B. (1999). Event-related fMRI of the auditory cortex. *Neuroimage*, 10, 417–429.
- Brown, S., Ngan, E., & Liotti, M. (2008). A larynx area in the human motor cortex. *Cerebral Cortex*, 18, 837–845.
- Calvert, G. A., Hansen, P. C., Iversen, S. D., & Brammer, M. J. (2001). Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the BOLD effect. *Neuroimage*, 14, 427–438.
- 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. *Journal of Neuroscience*, 18, 7426–7435.
- Crowder, R. G. (1989). Imagery for musical timbre. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 472–478.
- Cui, X., Jeter, C. B., Yang, D., Montague, P. R., & Eagleman, D. M. (2007). Vividness of mental imagery: Individual variability can be measured objectively. *Vision Research*, 47, 474–478.
- Daselaar, S. M., Porat, Y., Huijbers, W., & Pennartz, C. M. (2010). Modality-specific and modality-independent components of the human imagery system. *Neuroimage*, 52, 677–685.
- D'Ausilio, A., Altenmüller, E., Olivetti Belardinelli, M., & Lotze, M. (2006). Cross-modal plasticity of the motor cortex while listening to a rehearsed musical piece. *European Journal of Neuroscience*, 24, 955–958.
- Deiber, M. P., Ibañez, V., Sadato, N., & Hallett, M. (1996). Cerebral structures participating in motor preparation in humans: A positron emission tomography study. *Journal of Neurophysiology*, 75, 233–247.
- Dobson, M., & Markham, R. (1993). Imagery ability and source monitoring: Implications for eyewitness memory. *British Journal of Psychology*, 84, 111–118.
- Eberman, C., & McKelvie, S. J. (2002). Vividness of visual imagery and source memory for audio and text. *Applied Cognitive Psychology*, 16, 87–95.
- Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., & Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage*, 6, 218–229.
- Gonsalves, B., Reber, P. J., Gitelman, D. R., Parrish, T. B., Mesulam, M.-M., & Paller, K. A. (2004). Neural evidence that vivid imagining can lead to false remembering. *Psychological Science*, 15, 655–660.
- Griffiths, T. D. (2000). Musical hallucinosis in acquired deafness. Phenomenology and brain substrate. *Brain*, 123, 2065–2076.
- Grodd, W., Hülsmann, E., Lotze, M., Wildgruber, D., & Erb, M. (2001). Sensorimotor mapping of the human cerebellum: fMRI evidence of somatotopic organization. *Human Brain Mapping*, 13, 55–73.
- Gusnard, D. A., Akbudak, E., Shulman, G. L., & Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: Relation to a default mode of brain function. *Proceedings of the National Academy of Sciences, U.S.A.*, 98, 4259–4264.
- Halpern, A. R. (1988). Mental scanning in auditory imagery for songs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14, 434–443.
- Halpern, A. R. (1989). Memory for the absolute pitch of familiar songs. *Memory & Cognition*, 17, 572–581.
- Halpern, A. R., & Zatorre, R. J. (1999). When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies. *Cerebral Cortex*, 9, 697–704.
- Halpern, A. R., Zatorre, R. J., Bouffard, M., & Johnson, J. A. (2004). Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia*, 42, 1281–1292.
- Henson, R. N. A., Shallice, T., & Dolan, R. J. (1999). Right prefrontal cortex and episodic memory retrieval: A functional MRI test of the monitoring hypothesis. *Brain*, 122, 1367–1381.
- Herholz, S. C., Lappe, C., Knief, A., & Pantev, C. (2008). Neural basis of music imagery and the effect of musical expertise. *European Journal of Neuroscience*, 28, 2352–2360.
- Huijbers, W., Pennartz, C. M. A., Rubin, D. C., & Daselaar, S. M. (2011). Imagery and retrieval of auditory and visual information: Neural correlates of successful and unsuccessful performance. *Neuropsychologia*, 49, 1730–1740.
- Hyman, I. E., & Pentland, J. (1996). The role of mental imagery in the creation of false childhood memories. *Journal of Memory and Language*, 35, 101–117.
- Janata, P. (2001). Brain electrical activity evoked by mental formation of auditory expectations and images. *Brain Topography*, 13, 169–193.
- Janata, P., & Paroo, K. (2006). Acuity of auditory images in pitch and time. *Perception and Psychophysics*, 68, 829–844.
- Kanai, R., & Rees, G. (2011). The structural basis of inter-individual differences in human behaviour and cognition. *Nature Reviews Neuroscience*, 12, 231–242.
- 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., Schulze, K., Sammler, D., Fritz, T., Müller, K., & Gruber, O. (2009). Functional architecture of verbal and tonal working memory: An fMRI study. *Human Brain Mapping*, 30, 859–873.
- Kraemer, D. J., Macrae, C. N., Green, A. E., & Kelley, W. M. (2005). Musical imagery: Sound of silence activates auditory cortex. *Nature*, 434, 158.
- Lahav, A., Saltzman, E., & Schlaug, G. (2007). Action representation of sound: Audiomotor recognition network while listening to newly acquired actions. *Journal of Neuroscience*, 27, 308–314.
- Lawrie, S. M., Buechel, C., Whalley, H. C., Frith, C. D., Friston, K. J., & Johnstone, E. C. (2002). Reduced frontotemporal functional connectivity in schizophrenia associated with auditory hallucinations. *Biological Psychiatry*, 51, 1008–1011.
- Leaver, A. M., Van Lare, J., Zielinski, B., Halpern, A. R., & Rauschecker, J. P. (2009). Brain activation during anticipation of sound sequences. *Journal of Neuroscience*, 29, 2477–2485.
- Lewis, J. W., Beauchamp, M. S., & DeYoe, E. A. (2000). A comparison of visual and auditory motion processing in human cerebral cortex. *Cerebral Cortex*, 10, 873–888.
- Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory*, 12, 361–366.
- Lotze, M., Scheler, G., Tan, H. R., Braun, C., & Birbaumer, N. (2003). The musician's brain: Functional imaging of amateurs and professionals during performance and imagery. *Neuroimage*, 20, 1817–1829.
- Loucks, T. M. J., Poletto, C. J., Simonyan, K., Reynolds, C. L., & Ludlow, C. L. (2007). Human brain activation during phonation and exhalation: Common volitional control for two upper airway functions. *Neuroimage*, 36, 131–143.
- Macrae, C. N., Moran, J. M., Heatherton, T. F., Banfield, J. F., & Kelley, W. M. (2004). Medial prefrontal activity predicts memory for self. *Cerebral Cortex*, 14, 647–654.
- Mechelli, A., Price, C. J., Friston, K. J., & Ishai, A. (2004). Where bottom-up meets top-down: Neuronal interactions during perception and imagery. *Cerebral Cortex*, 14, 1256–1265.
- Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, 6, 688–691.
- Petrides, M. (2000). The role of the mid-dorsolateral prefrontal cortex in working memory. *Experimental Brain Research*, 133, 44–54.
- Peynircioglu, Z. F., Tekcan, A. I., Wagner, J. L., Baxter, T. L., & Shaffer, S. D. (1998). Name or hum that tune: Feeling of knowing for music. *Memory & Cognition*, 26, 1131–1137.
- Phan, K. L., Taylor, S. F., Welsh, R. C., Decker, L. R., Noll, D. C., Nichols, T. E., et al. (2003). Activation of the medial prefrontal cortex and extended amygdala by individual ratings of emotional arousal: A fMRI study. *Biological Psychiatry*, 53, 211–215.
- Platel, H., Baron, J. C., Desgranges, B., Bernard, F., & Eustache, F. (2003). Semantic and episodic memory of music are subserved by distinct neural networks. *Neuroimage*, 20, 244–256.
- Richter, W., Andersen, P. M., Georgopoulos, A. P., & Kim, S. G. (1997). Sequential activity in human motor areas during a

- delayed cued finger movement task studied by time-resolved fMRI. *NeuroReport*, 8, 1257–1261.
- Schacter, D. L., Buckner, R. L., Koutstaal, W., Dale, A. M., & Rosen, B. R. (1997). Late onset of anterior prefrontal activity during true and false recognition: An event-related fMRI study. *Neuroimage*, 6, 259–269.
- Schürmann, M., Raji, T., Fujiki, N., & Hari, R. (2002). Mind's ear in a musician: Where and when in the brain. *Neuroimage*, 16, 434–440.
- Serafine, M. L., Davidson, J., Crowder, R. G., & Repp, B. H. (1986). On the nature of melody-text integration in memory for songs. *Journal of Memory and Language*, 25, 123–135.
- Wallace, W. T. (1994). Memory for music: Effect of melody on recall of text. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 1471–1485.
- Watanabe, T., Yagishita, S., & Kikyo, H. (2008). Memory of music: Roles of right hippocampus and left inferior frontal gyrus. *Neuroimage*, 39, 483–491.
- Wheeler, M. E., Petersen, S. E., & Buckner, R. L. (2000). Memory's echo: Vivid remembering reactivates sensory-specific cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, 97, 11125–11129.
- White, K., Ashton, R., & Brown, R. (1977). The measurement of imagery vividness: Normative data and their relationship to sex, age, and modality differences. *British Journal of Psychology*, 68, 203–211.
- Yoo, S. S., Lee, C. U., & Choi, B. G. (2001). Human brain mapping of auditory imagery: Event-related functional MRI study. *NeuroReport*, 12, 3045–3049.
- Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences*, 6, 37–46.
- Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: Auditory–motor interactions in music perception and production. *Nature Reviews Neuroscience*, 8, 547–558.
- Zatorre, R. J., Evans, A. C., & Meyer, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. *Journal of Neuroscience*, 14, 1908–1919.
- Zatorre, R. J., & Halpern, A. R. (2005). Mental concerts: Musical imagery and auditory cortex. *Neuron*, 47, 9–12.
- Zatorre, R. J., Halpern, A. R., & Bouffard, M. (2010). Mental reversal of imagined melodies: A role for the posterior parietal cortex. *Journal of Cognitive Neuroscience*, 22, 775–789.
- Zatorre, R. 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. *Journal of Cognitive Neuroscience*, 8, 29–46.