

When That Tune Runs Through Your Head: A PET Investigation of Auditory Imagery for Familiar Melodies

Andrea R. Halpern and Robert J. Zatorre¹

Psychology Department, Bucknell University, Lewisburg, PA 17837, USA and ¹Montreal Neurological Institute, McGill University, Montreal, Canada

The present study used positron emission tomography (PET) to examine the cerebral activity pattern associated with auditory imagery for familiar tunes. Subjects either imagined the continuation of nonverbal tunes cued by their first few notes, listened to a short sequence of notes as a control task, or listened and then reimagined that short sequence. Subtraction of the activation in the control task from that in the real-tune imagery task revealed primarily right-sided activation in frontal and superior temporal regions, plus supplementary motor area (SMA). Isolating retrieval of the real tunes by subtracting activation in the reimagine task from that in the real-tune imagery task revealed activation primarily in right frontal areas and right superior temporal gyrus. Subtraction of activation in the control condition from that in the reimagine condition, intended to capture imagery of unfamiliar sequences, revealed activation in SMA, plus some left frontal regions. We conclude that areas of right auditory association cortex, together with right and left frontal cortices, are implicated in imagery for familiar tunes, in accord with previous behavioral, lesion and PET data. Retrieval from musical semantic memory is mediated by structures in the right frontal lobe, in contrast to results from previous studies implicating left frontal areas for all semantic retrieval. The SMA seems to be involved specifically in image generation, implicating a motor code in this process.

Introduction

Cognitive scientists are interested in the mental structures that underlie the experience of imagery, or mental acts in which we seem to re-enact the experience of perceiving an object when the object is no longer available. Cognitive psychologists have wondered whether this experience is fundamentally different to the more abstract mentation used in recalling facts or solving arithmetic problems. Purely behavioral methods have shown intriguing similarities between performance on perceptual and imaginal versions of the same task (Farah, 1989) or facilitation by an imagined stimulus on performance of a task involving a perceived stimulus (Hubbard and Stoeckig, 1988). However, these behavioral methods have their limitations in helping us decide whether perception and imagery share similar mental structures. For instance, although similarity of response patterns in perceived and imagined tasks may indicate shared mental structures, the similarity may be coincidental or epiphenomenal.

To investigate the nature of mental imagery further, researchers have turned to physiological evidence that imagery and perception may share actual neural structures. Farah reviewed evidence from brain-damaged patients who show parallel deficits in visual imagery and perception skills after damage to particular brain areas (Farah, 1988). More directly, a number of researchers have employed brain-imaging technology to observe the brain areas that are active while participants perceive or imagine stimuli. To the extent that brain areas known to be associated with sensory processing are active during imagery tasks, we may conclude that the brain efficiently

uses similar areas both to process information initially, as well as to reactivate it for further processing.

For the case of visual imagery, several studies have indeed found the hypothesized activation in visual cortical areas during imagery tasks (Kosslyn *et al.*, 1993) [reviewed by Farah (Farah, 1995) and Mellet *et al.* (Mellet *et al.*, 1998)]. Many studies investigating visual imagery have found evidence that visual association areas (and sometimes primary areas) are engaged during visual imagery tasks. This pattern obtains over different kinds of imagery tasks and different brain-imaging techniques, including SPECT (Goldenberg *et al.*, 1989) and ERP (Farah *et al.*, 1989) as well as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) mentioned above.

In addition to the localization of visual areas active in imagery, researchers have investigated the lateralization of imagery processes. The findings related to this question are mixed. But as Mellet and co-workers have pointed out, the degree of lateralization may depend on both the complexity of the task (leading to more right-sided activation) and nameability of the stimuli (leading to more left-sided activation) (Mellet *et al.*, 1998).

All these studies have examined visual imagery exclusively. To what extent can we extend conclusions made in the visual domain to other domains, specifically audition? Behavioral investigations of auditory imagery have suggested that it can be manipulated and measured in ways similar to that of visual imagery. For instance, Halpern asked people to mentally compare the pitches corresponding to two words drawn from familiar tunes (Halpern, 1988). She found that the time taken to respond was systematically related to the number of beats separating the words in the real tune. The interpretation here was that as visual images represent space, auditory images may represent time or other auditory properties, such as loudness (Farah and Smith, 1983). This certainly accords with subjective reports that people can 'hear' sounds in their heads.

In work previous to the current investigation, we asked whether auditory imagery for music might be mediated by the same brain areas as auditory perception. Our first study (Zatorre and Halpern, 1993) presented the mental pitch comparison task described above to patients who had undergone right or left temporal-lobe excisions for the relief of epilepsy, plus normal controls. We also presented a perceptual version in which the pitches were to be compared while the song in question was actually being presented. Our results were straightforward: all groups found the imagery task more difficult, as expected, but only the right temporal lobectomy group showed a performance deficit. They were impaired relative to the other groups by the same amount on both imagery and perception tasks.

This result seemed to implicate the right temporal lobe as being necessary to perform both auditory imagery and perception tasks, but by this approach we could not see what structures would normally be active during performance of such a task.

Consequently, our next study (Zatorre *et al.*, 1996) utilized PET to look at cerebral blood flow (CBF) in normal volunteers during essentially the same imagery and perception tasks described above. Each of these experimental conditions was subtracted from a visual baseline, in which the words from the songs were randomly paired and presented on a screen for a visual length judgment. The subtractions revealed remarkable similarity in CBF patterns in the perception and imagery conditions. Notably, both tasks activated auditory association regions in the superior temporal gyrus (STG) bilaterally, as well as several areas bilaterally in the frontal lobe and one parietal area. The supplementary motor area (SMA) was also activated in both tasks. Outside of activation unique to primary auditory cortex in the perception condition (the only one with actual auditory input), only four brain regions showed statistically significant differences between imagery and perception tasks, including frontopolar areas bilaterally, the subcallosal gyrus and, in the only lateralized effect we found, right thalamus.

These two experiments confirmed the importance of auditory association areas in mediating auditory imagery, analogously to the visual imagery results described previously. However, we were left with several open questions. First, it is evident that our imagery and perception tasks were complex. They involved retrieval of a tune from musical semantic memory (in the imagery task), the rehearsal of the first pitch while the second was retrieved (a working memory task in both), and finally the pitch comparison and decision. We could not separate these components and thus link them to various areas in the frontal lobes that have been implicated in the literature on brain areas active in memory tasks. Second, the two experiments seemed to be at variance with one another concerning lateralized effects. The lesion study suggested that the right temporal lobe is essential for processing heard and imagined musical stimuli. This is consistent with numerous other studies showing an important role for the right temporal neocortex in tonal perception tasks (Milner, 1962; Zatorre, 1985, 1988; Divenyi and Robinson, 1989; Robin *et al.*, 1990; Zatorre and Samson, 1991; Zatorre *et al.*, 1994; Samson and Zatorre, 1994; Liégeois-Chauvel *et al.*, 1998). However, the PET study of musical imagery showed a bilateral pattern of activation in the STG.

Regarding the first set of concerns, several groups of researchers have been investigating the involvement of the frontal lobes in working memory tasks. For example, Petrides *et al.* have found using PET that the dorsolateral frontal cortex bilaterally (Brodmann areas 46 and 9) is important in mediating self-generated working memory tasks, such as randomly generating numbers from 1 to 10 without repeating any (Petrides *et al.*, 1993). Smith and Jonides have also reported dorsolateral frontal activity in a working memory task in which subjects have to compare the current stimulus with the stimulus two or three positions back in a list for a same-different judgment (Smith and Jonides, 1997). Braver *et al.* parametrically varied working memory load (from zero-back to three-back comparisons) and found that activation in areas 46 and 9 increased monotonically with increasing working memory load (Braver *et al.*, 1997).

The other major component of our original imagery task – retrieval from semantic memory – has been somewhat less studied. Nyberg *et al.* have reviewed the possible differences in brain areas active during retrieval from episodic and semantic memory (Nyberg *et al.*, 1996). In their HERA model (hemispheric encoding/retrieval asymmetry), they implicate the left prefrontal area for semantic retrieval (which they also identify with episodic encoding, on the assumption that retrieval of a

semantic fact will be newly encoded as an episodic memory). In contrast, they implicate the right prefrontal area in episodic retrieval. Gabrieli *et al.* similarly found that judging words as being abstract or concrete – a semantic memory task – selectively activated the left inferior frontal gyrus (areas 45, 46 and 47), as measured by fMRI (Gabrieli *et al.*, 1996). However, the role of these regions of the left inferior frontal cortex may be more general, since numerous studies have consistently found CBF increases in this location during tasks that require lexical search and retrieval, including noun-verb generation (Petersen *et al.*, 1988), synonym generation or translation (Klein *et al.*, 1995), or stem completion (Buckner *et al.*, 1995).

Most of these studies have used verbal stimuli. Petrides and colleagues usually find bilateral activation in their verbal working memory tasks, although Smith and Jonides have evidence that verbal working memory is left lateralized whereas spatial working memory is right lateralized (Smith *et al.*, 1996). Tulving and colleagues, however, assert that the episodic/semantic distinction in retrieval cuts across materials. Nyberg *et al.* cite numerous cases of object and face memory that seem to follow the left = semantic and right = episodic retrieval scheme (Nyberg *et al.*, 1996). None of the studies in this literature have used music, which, as noted above, is known to be heavily dependent on right hemisphere structures. This brings us to the second concern remaining from our prior PET study, the lack of asymmetrical right-sided activations in either the imagery or perception task. Recall, however, that in both of our previous studies the music we used had lyrics associated with them. It is possible that the bilateral findings of our first PET study were due to the fact that the task required processing of both words and music.

In view of both the memory literature cited and our own previous research, we decided to construct a musical imagery task that did not require processing of lyrics. By presenting stimuli that were exclusively musical, we minimized the involvement of verbal processing structures in the left hemisphere. This would, we hoped, reveal music-specific structures in the right hemisphere in the temporal lobe (as predicted from both our previous studies), as well as the frontal lobe. In addition, we modified our task to try to better separate the retrieval of information from musical semantic memory versus the working memory task of keeping the memory traces active for a period of time. The main imagery task (Cue/Imagery) involved presentation of the first few notes of familiar, but nonverbal tunes, such as movie themes and excerpts from classical music. Participants then had to imagine the rest of the tune, and press a button when they completed the task. This involved imagery and working memory processes, as well as retrieval of the tune from semantic memory. In the Control task, a novel tone sequence derived from the real tune fragment was presented, and subjects simply pressed a button at the end of each one. In a second control task called Control/Imagery, the same novel tone sequence was presented and subjects had to reimagine the sequence, then press a button. This task involves working memory (because the sequence had to be remembered) and imagery (because it had to be rehearsed), but does not require retrieval from semantic memory, as the sequence was novel. Subtracting the Control task from the Control/Imagery task should remove the effects of hearing a note sequence and thereby isolate imagery and working memory processes. Subtracting the Control/Imagery task from the Cue/Imagery task should remove the effects of auditory input, working memory and imagery, thereby isolating

Table 1

Summary of experimental paradigm

Condition	Stimulus	Task	Imagery required?	Retrieval required?
Control	control tone sequence	listen	no	no
Cue/Image	cue sequence	listen and image rest of tune	yes	yes
Control/Image	control tone sequence	listen and image control sequence	yes	no

musical semantic retrieval. A summary of these tasks is presented in Table 1.

To summarize, we had four goals for this experiment. First, we were interested in testing the generality of our results from our first PET study (Zatorre *et al.*, 1996) with new sets of materials and a new task. Specifically, we hoped to confirm the activation of the auditory association areas in the STG during silent auditory imagery tasks. We also sought to confirm the activation of SMA and frontal areas during our tasks. Second, we hypothesized that eliminating all words from our materials would show asymmetrical activation in right temporal and frontal areas. Third, we were interested in seeing whether working memory associated with musical tasks would activate areas in the dorsolateral frontal lobe as found by previous researchers. Although most previous work has found bilateral or left-sided activation during verbal working memory tasks, Smith *et al.* have suggested that spatial working memory is mediated by right prefrontal structures (Smith *et al.*, 1996). Music provides an interesting test of the generality of this finding. Thus we were particularly interested to see if we found any asymmetry in dorsolateral frontal cortex activation when listeners had to imagine a novel sequence just presented. We also wondered whether SMA activation would be associated with the working memory aspect of our task. Our thinking here was that activation of motor codes, with which the SMA has been associated (Rao *et al.*, 1997), may assist listeners in preserving memory traces of tunes.

Finally, the aspect of our task requiring retrieval from musical semantic memory provides a test of the HERA model of Tulving and colleagues, in which semantic retrieval is associated exclusively with left frontal areas. If our task elicits substantial right frontal activation when the familiar tune fragments are presented for mental continuation, we would have to question the generality of the HERA formulation in favor of a material-specific scheme.

Materials and Methods

Subjects

Eight right-handed volunteers (five women, three men, mean age 24) participated after giving informed consent in accord with ethical guidelines in place at the Montreal Neurological Institute. Subjects had received varying amounts of musical training, with a range of 3–16 years of formal music lessons, and an average of 9.5 years.

Stimuli

Three types of stimuli were prepared: melodic themes, cue sequences and control sequences (Fig. 1). The melodic themes were only used for familiarization; during scanning only cue and control sequences were used. Melodies were initially selected based on two criteria: that they not be associated with lyrics, and that they were rated as familiar in pilot testing. Fifteen melodies were chosen from among classical music (e.g. dances from the *Nutcracker Suite*, opening theme from Beethoven's Fifth Symphony), television shows or movies (e.g. themes from *Dallas*, *Star Wars*), and from other popular sources (e.g. chimes of Big Ben, Scott Joplin's 'The Entertainer'). The initial few bars of each of these melodies

Theme from "Dallas"



Cue sequence



(imagine continuation)

Control sequence



Figure 1. Illustration of the stimuli used, in musical notation. The first line illustrates a melodic theme taken from the television show *Dallas* which was played to subjects during screening sessions and prior to scanning. The second line illustrates the cue sequence for this item (consisting of the first five notes of the theme), which was used during the Cue/Image condition. Subjects were instructed to imagine the rest of the theme continuing from the cue sequence. The third line shows a control sequence (consisting of a random permutation of the five notes), which was used during the Control and Control/Image conditions.

were then selected as the melodic themes to be used for the study. One additional feature of these themes is that they were selected to fall into one of three categories based on their duration, with average durations of short, medium and long themes being 2.2, 4.8 and 6.2 s respectively. This manipulation enabled us to measure the time taken to imagine the theme. The end point of each theme coincided with phrasal boundaries.

Fifteen cue sequences were then created by taking the first few notes from each theme (Fig. 1). Pilot testing ensured that the cues uniquely specified one of the target melodies, and that subjects had no difficulty generating an internal image from each cue. Finally, 15 control sequences were created by randomly permuting the tones within each cue sequence, to create a set of control sequences that were matched for number of tones, total duration and types of rhythmic and tonal intervals.

Procedure

Subjects were screened to ensure familiarity with the tunes to be used, and to indicate to them the endpoints selected for each melody. Subjects rated each melody for familiarity; all subjects who were retained for scanning rated the stimuli as familiar or very familiar (average rating of 1.2 on a scale of 1–5). Subjects were also instructed to pay close attention to the endpoint of each melody. On the day of scanning, subjects were once again presented with each of the themes to remind them of the stimulus materials to be used, and to help them recall the endpoints.

Three conditions were tested during scanning: Control, Cue/Image and Control/Image, in that order (see Table 1). In the Control condition, subjects were presented with each of the control sequences described above; they were instructed simply to listen to each sequence, and to press a mouse key after each stimulus. In the Cue/Image condition, subjects heard the cues associated with the themes (i.e. the first few notes of the tune), and were asked to imagine the continuation of each melody as it followed from the cue. They were further instructed to stop imagining the melody at the same point as had been demonstrated during screening, and to press the mouse key at that point. In the Control/Image condition subjects were told to listen to the control sequences, and were then instructed to imagine the same sequence just heard, and to press the key when this was accomplished.

Practice trials were given prior to each scan condition using the same stimuli but in a different randomization. Intertrial intervals were 5, 6 or 7 s for the short, medium and long trials respectively (to allow sufficient

Table 2

Stereotaxic coordinates and significance levels of activation foci in the Cue/Image – Control subtraction. The two far right columns indicate foci in corresponding locations for the Cue/Image – Control subtraction (Table 3) and the Control/Image – Control subtraction (Table 4)

Region	x	y	z	t	Corresponding foci in other conditions	
					Table 3	Table 4
<i>Right frontal cortex</i>						
1. Inferior frontal gyrus (10/47)	34	53	-9	4.57	1	
	38	42	-3	4.13		
2. Inferior frontal gyrus (45)	46	13	6	6.22	2	
3. Middle frontal gyrus (46)	36	46	23	3.63	3	
4. Precentral gyrus (6)	51	-1	47	4.20		
<i>Left frontal cortex</i>						
5. Inferior frontal gyrus (45)	-42	12	2	3.70	4	
	-35	22	5	4.28		
6. Middle frontal gyrus (46/9)	-34	44	18	3.84	5	
7. Middle frontal gyrus (44)	-48	10	23	5.49		3
8. Precentral gyrus (6)	-48	-6	41	4.78		6
<i>Right temporal cortex</i>						
9. Superior temporal gyrus (22)	56	-30	8	3.41	6	
10. Inferior temporal gyrus (37)	62	-42	-8	4.40	7	
<i>Other regions</i>						
11. Supplementary motor area (6)	-1	8	62	7.47		7
12. Anterior cingulate (32)	-4	29	30	3.53	8	
13. Right parietal cortex (40)	46	-49	51	4.87	9	
14. Precuneus (7)	-4	-69	36	3.64		

time for subjects to image the appropriate duration for each trial). The tasks were begun prior to the onset of scanning, which typically started between the third and fifth trials. Stimuli were presented binaurally over insert earphones (EAR Tone Type 3A), which had been calibrated for an average intensity of 72 dB SPL (A).

PET Scanning

PET scans were obtained with a Siemens Exact HR+ tomograph operating in three-dimensional acquisition mode. The distribution of CBF was measured during each 60 s scan using the H₂¹⁵O water bolus method (Raichle *et al.*, 1983). MRI scans (160 1 mm thick slices) were also obtained for each subject with a 1.5 T Phillips ACS system to provide anatomical detail. CBF images were reconstructed using a 14 mm Hanning filter, normalized for differences in global CBF, and coregistered with the individual MRI data (Evans *et al.*, 1992). Each matched MRI/PET data set was then linearly resampled into the standardized stereotaxic coordinate system of Talairach and Tournoux (Talairach and Tournoux, 1988) via an automated feature-matching algorithm (Collins *et al.*, 1994). PET images were averaged across subjects for each condition, and the mean change image volume obtained for each comparison; this volume was converted to a *t*-statistic map, and the significance of focal CBF changes was assessed by a method based on three-dimensional Gaussian random-field theory (Worsley *et al.*, 1992). The presence of significant changes in CBF was first established on the basis of an exploratory search, for which the *t*-value criterion was set at 3.53 or greater. This value corresponds to an uncorrected *P*-value of 0.0004 (two-tailed), and results in an average of 0.58 false positives per search volume of 182 resolution elements (dimensions of 14 × 14 × 14 mm), corresponding approximately to the volume of gray matter scanned. For the superior temporal region, where activity had been predicted based on previous findings, the threshold was lowered to *t* = 3.0.

Results

Behavioral Data

All subjects indicated that they had been able to generate the musical images required during the Cue/Image and Control/

Image tasks, and subjectively felt that they had a strong sense of vivid auditory imagery ('hearing the music in my head'). Mean latencies for seven of the eight subjects (data for the eighth subject were lost due to computer error) to key press, measured from the onset of the cue in the Cue/Image condition for the short, medium and long trials were 2.72, 3.72 and 4.55 s respectively. These values were entered into an analysis of variance, which indicated a significant difference among them [$F(2,12) = 32.36, P < 0.001$]; this provides behavioral evidence that subjects were generating an auditory image in conformity with the desired stimulus duration. All seven subjects showed the pattern of increasing latency with increasing length of theme.

Analysis of CBF data

Comparisons were performed by subtracting the Control task from each of the other two, and also by comparing the Cue/Image and the Control/Image tasks to one another. Table 2 shows the stereotaxic coordinates and *t*-values for foci in the Cue/Image – Control subtraction. In addition, the two rightmost columns show the foci from the other two subtractions that correspond in location to those in Table 2.

The Cue/Image – Control subtraction was intended to capture all the processes involved in musical imagery, controlling for physical stimulus input and response output. One question of interest was whether greater activity would be detected in the right frontal region than in the left in this subtraction. The Cue/Image – Control comparison did yield significant activation within the right inferior frontal gyrus (focus 1, area 10/47), with no homologous activation on the left. A second focus in right frontal cortex (focus 2, area 45; see Fig. 2, top), was matched by a similar region of activity on the left (focus 5), but at a much lower level of significance. A third region in the middle frontal gyrus was approximately equally active in the two hemispheres (foci 3 and 6, area 46). The only region that was uniquely active in the left frontal lobe was located approximately within area 44 (focus 7).

In addition, significant CBF increases were noted in the predicted area of the right superior temporal cortex (focus 9, visible in Fig. 2, top panel; the *t*-value of 3.41 in this case falls just below the *t*-threshold for exploratory search set *a priori*, but is well above the threshold for predicted activation sites), and also in the right inferior temporal cortex (focus 10). No significant activity was detected in the left temporal lobe, even using the lower *t*-threshold value. Finally, and also in keeping with the predictions, this subtraction yielded activity within the SMA (focus 11; Fig. 2).

The Cue/Image – Control/Image subtraction (Table 3 and Fig. 2, middle) was intended to isolate processing components associated with retrieving real tunes from semantic memory. This comparison yielded a number of foci that were very similar in location to those elicited by the Control/Image – Control subtraction. In fact, all nine foci in Table 3 are within regions also shown in Table 2, as indicated. The inferior frontal gyrus, area 10/47 (focus 1), again showed a significant CBF increase in the right hemisphere only, as in the previous subtraction. Areas 45 and 46 (foci 2–5) showed bilateral activation, but with much higher activity on the right. Importantly, this subtraction also yielded significant activity within the right STG (focus 6), as well as the right inferior temporal gyrus (focus 7), without any corresponding activation on the left. It is notable, however, that this comparison did not yield any detectable activity within the SMA, even applying the less stringent statistical cutoff for predicted areas.

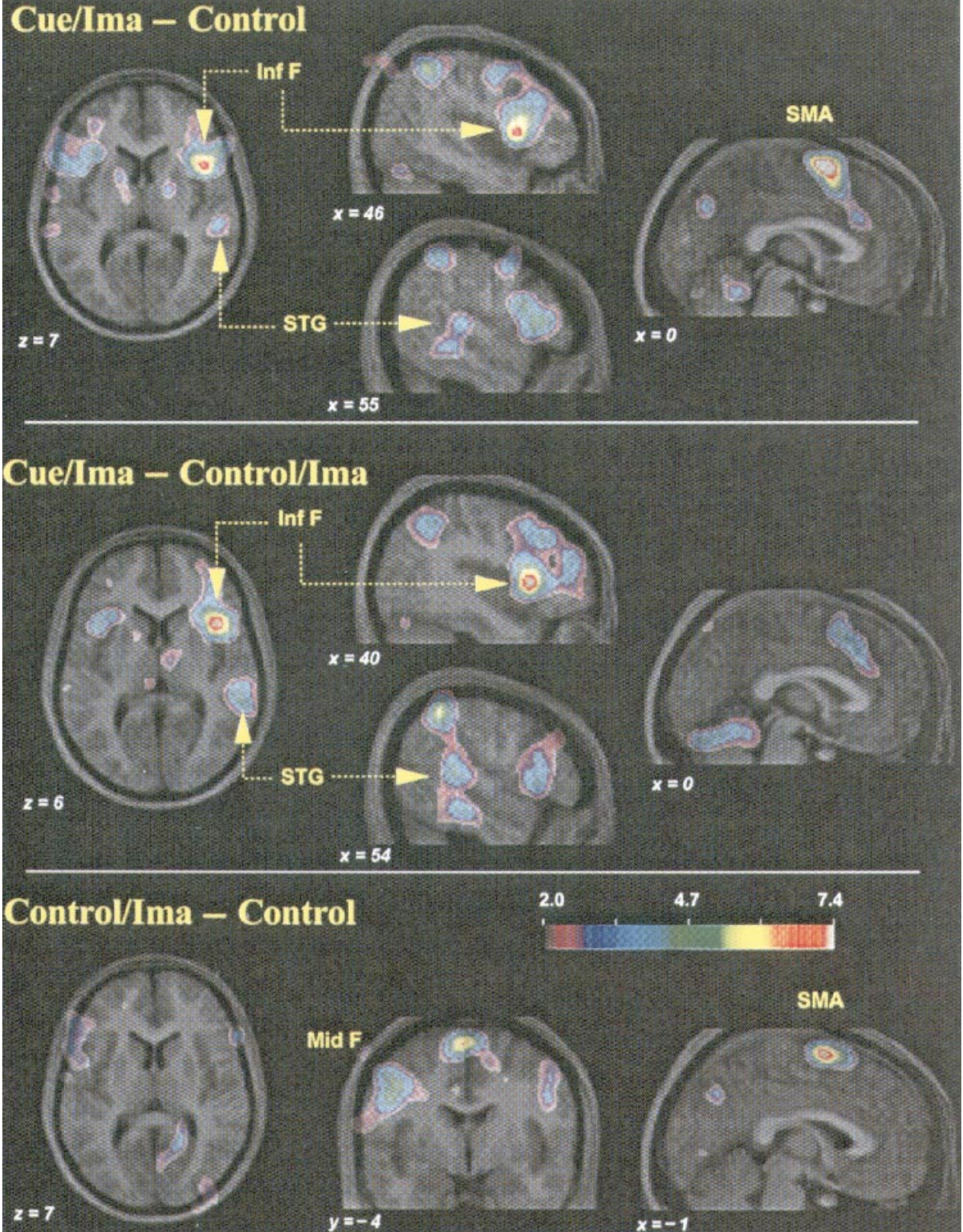


Figure 2. Merged PET/MRI images illustrating selected regions of significant increase in CBF in each of three comparisons. (Top panel) Areas of CBF increase in the Cue/Image – Control subtraction (keyed to Table 2). The leftmost image corresponds to a horizontal section ($z = 7$) and shows the activity in the right and left inferior frontal gyrus, corresponding to foci 2 and 5 respectively in Table 2. Also visible in the horizontal section is activation in the right superior temporal gyrus (focus 9 in Table 2). The right frontal and temporal areas of activity are also shown in the two parasagittal sections through the right hemisphere shown in the middle of the top panel ($x = 46$ and 55). The positions of the dotted lines indicate corresponding planes of section in the figures. The rightmost image in the top panel shows the SMA activity in a midsagittal section (focus 11 in Table 2). (Middle panel) Areas of CBF increase in the Cue/Image – Control/Image subtraction (keyed to Table 3). The horizontal section ($z = 6$) and the two parasagittal sections ($x = 40$ and 54) illustrate the activation in right and left frontal areas (foci 2 and 4 of Table 3), and in the right superior temporal cortex (focus 6). Note the similarity of activation in right and left inferior frontal areas and right superior temporal cortex to the upper panel. Note also the absence of SMA activity in the midsagittal section (far right). (Bottom panel) Images associated with the Control/Image – Control subtraction (keyed to Table 4). The horizontal section ($z = 7$) is included to demonstrate the absence of activation in frontal or temporal areas comparable to those shown in the other two panels. The middle image is a coronal section ($y = 4$) to show the region of left midfrontal activity (focus 3 in Table 4). Also visible in this section and in the midsagittal section (far right) is activity within the SMA (focus 7 in Table 4).

Table 3

Stereotaxic coordinates and significance levels of activation foci in the Cue/Image – Control/Image subtraction

Region	x	y	z	t
<i>Right frontal cortex</i>				
1. Inferior frontal gyrus (10/47)	29	55	-8	4.28
2. Inferior frontal gyrus (45)	40	17	6	7.41
3. Middle frontal gyrus (46)	36	46	23	4.29
<i>Left frontal cortex</i>				
4. Inferior frontal gyrus (45)	-42	15	2	4.65
5. Middle frontal gyrus (46/9)	-32	46	26	3.91
<i>Right temporal cortex</i>				
6. Superior temporal gyrus (22)	55	-40	11	4.53
	61	-42	3	4.40
7. Inferior temporal gyrus (37)	60	-41	-11	4.61
<i>Other regions</i>				
8. Anterior cingulate (32)	4	24	41	4.48
9. Right parietal cortex (40)	48	-49	50	6.31

The final comparison, Control/Image – Control (Table 4; Fig. 2, bottom), was designed to demonstrate the activity associated with imagery in the absence of any semantic components. This subtraction did not show activation in the inferior frontal areas that were detected in the other subtractions, nor in the temporal cortex, even with a less stringent statistical threshold value. It did, however, demonstrate a clear CBF increase in the SMA (similar in location to that shown in Table 2), along with predominantly left-sided frontal cortical sites. The left dorsolateral frontal region, area 44 (focus 3), was similar in location to that seen in Table 2.

Discussion

The principal findings of this study confirm predictions that activity in the right auditory association cortex, together with the SMA, accompanies musical imagery (Cue/Image – Control; see Fig. 2). Breaking the task down into its components, we found that when imagery entails retrieval from musical semantic memory (Cue/Image – Control/Image), activity ensued in a right inferior frontal region and bilaterally in middle frontal areas (more significant on the right side), together with right auditory association areas in STG. When imagery does not require semantic retrieval (Control/Image – Control), left frontal areas and SMA are recruited.

Auditory Areas Active in Imagery

The present study extends previous findings implicating auditory cortical regions in musical imagery (Zatorre and Halpern, 1993; Zatorre *et al.*, 1996). Even using a different behavioral paradigm and stimuli from the previous studies, several commonalities emerge. The most salient point is that auditory association areas are involved in processing imagined familiar melodies. Because the task design involved similar auditory input in each scan condition, the activation in auditory cortex during imagery must be due to processing beyond that elicited by the auditory stimulation. This pattern of activation supports the hypothesis that cortical perceptual areas can mediate internally generated information. This conclusion is consistent with findings from the visual domain (Kosslyn *et al.*, 1993; Farah, 1995).

Also consistent with prior PET data (Zatorre *et al.*, 1996), only associative cortical regions, not primary, were active in the

Table 4

Stereotaxic coordinates and significance levels of activation foci in the Control/Image – Control subtraction

Region	x	y	z	t
<i>Right frontal cortex</i>				
1. Precentral gyrus (6/4)	26	-30	53	3.89
<i>Left frontal cortex</i>				
2. Inferior frontal gyrus (47)	-48	30	-17	3.81
3. Middle frontal gyrus (44)	-51	6	20	3.81
4. Frontal pole (10)	-17	53	26	3.54
5. Superior frontal gyrus (8)	-21	12	50	5.04
6. Precentral gyrus (6)	-46	-7	39	4.63
<i>Other regions</i>				
7. Supplementary motor area (6)	-1	1	65	6.95
8. Right lingual gyrus (19)	20	-54	6	3.73
9. Right superior occipital gyrus (19)	17	-83	33	3.73

imagery task. Comparing the locus of activity in the superior temporal region with anatomical probability maps derived from MR scans in stereotaxic space indicates that the activation is well posterior to Heschl's gyrus (Penhune *et al.*, 1996), and is located within or just posteroventral to the planum temporale (Westbury *et al.*, 1999). These areas are known from physiological and anatomical studies to constitute unimodal auditory association cortex (Celesia, 1976; Galaburda and Sanides, 1980). To date, therefore, auditory imagery paradigms for music, as well as for simpler tonal stimuli (Rao *et al.*, 1997; Penhune *et al.*, 1998) have not revealed activation in primary auditory cortex, in contrast to at least some visual imagery tasks which have found activation in primary visual areas (Kosslyn *et al.*, 1993). Whether this reflects differences related to how information is processed in different modalities, or to task design, or other factors remains to be determined.

Laterality of Effects

Another important point concerns the laterality of the STG activity. Our hypothesis was that nonverbal melodies would activate primarily right temporal cortex, in contrast to the bilateral activity observed previously with verbal melodies. The present finding of right auditory cortical activation therefore supports the broader hypothesis that mechanisms within the right hemisphere are specialized for processing tonal patterns, as predicted based on previous lesion (Milner, 1962; Zatorre and Samson, 1991; Liégeois-Chauvel *et al.*, 1998;) and imaging (Démonet *et al.*, 1994; Zatorre *et al.*, 1994; Binder *et al.*, 1997;) studies. Specifically, the present PET data are in good accord with the results of our behavioral lesion study (Zatorre and Halpern, 1993), in which we observed that right temporal-lobe resection resulted in decrements on perceptual and imaginal music tasks, whereas similar damage to the left temporal region had no effect.

An important conclusion to be drawn from the present results is that the right-hemisphere specialization extends beyond perceptual analysis to encompass complex tonal imagery processes. Supporting evidence for this conclusion comes from two studies in which subjects were asked to listen to a simple tonal sequence and then either continue tapping in the same rhythm (Rao *et al.*, 1997) or tap in imitation of the sequence (Penhune *et al.*, 1998). Both studies reported activation within right but not left posterior STG, which the authors interpreted as reflecting an auditory imagery process that accompanied the tapping. In

support of this, Janata found that the scalp topography of the electrical activity elicited by imaging the continuation of a melody is similar to the N100 component elicited by a real note (Janata, 1999).

Image Generation versus Retrieval

The experimental design of the present study allowed us to dissociate some of the processing components associated with retrieving and imagining a familiar tune. The Cue/Image - Control/Image subtraction isolated processes related to retrieval of the tune from semantic memory, whereas the Control/Image - Control subtraction focused on image generation without a retrieval component. These two subtractions revealed complementary areas of activation (see Table 2). The former comparison showed the right STG activation referred to above, together with predominantly right inferior frontal cortical activation. The latter subtraction did not show activation in these areas, but instead showed activity in SMA and in several left frontal regions. Taken together, these two subtractions yield an activity pattern similar to that seen in the Cue/Image - Control subtraction, suggesting that our design successfully captured the decomposition of the complex imagery task into separate retrieval and generation components.

In this context, the frontal-lobe activity in the Cue/Image - Control/Image subtraction may be interpreted as reflecting retrieval from musical semantic memory. Of the frontal regions activated (see Table 3), the most inferior one (area 10/47) was exclusively seen on the right, and the other two (areas 45 and 46) were active bilaterally but with a higher *t*-value on the right. The right inferior frontal focus observed in this subtraction is comparable to one found in our previous study (Zatorre *et al.*, 1996), in the comparison of imagery to perceptual conditions (the coordinates of the activity observed in that study, 34, 53, -11, are within 2 mm of focus 1 in Table 2). That subtraction was meant to isolate image retrieval and generation from perceptual processes. Although the paradigms in these studies were different, both had in common the necessity to retrieve a stored representation of a tune based on a cue. It is of interest to note that the STG areas identified in the present study are probably homologous to regions which in the macaque have been shown to be topographically interconnected with inferior frontal cortical areas (Petrides and Pandya, 1988; Romanski *et al.*, 1999). The CBF changes observed in the inferior frontal regions may therefore be interpreted as reflecting activity within this functional network.

Another region implicated in musical semantic retrieval in our earlier study was the right thalamus. In the current study a similar region was activated in the Cue/Image - Control/Image subtraction, albeit just below our relatively stringent level of significance for an exploratory search (coordinates: 8, -11, 8; *t* = 3.24). These convergent findings therefore further implicate a right inferior frontal/thalamic network in melodic semantic retrieval.

The frontal cortical areas just described are approximately homologous to the areas in the left hemisphere that have been proposed as mediating retrieval from verbal semantic memory in the HERA model (Nyberg *et al.*, 1996). We propose that the neural substrate of semantic memory retrieval may depend on the type of material to be retrieved. Retrieval of familiar musical information may involve the right hemisphere predominantly, as has already been established for the perception and discrimination of many musical materials. These findings indicate

the importance of using music to extend the generality of processing models.

An alternative interpretation of our findings concerning the retrieval component is that the imagery task may have entailed some degree of episodic retrieval. This could have occurred since subjects were asked to image the melodic theme to a specific endpoint, as demonstrated in screening and pre-scanning sessions. Thus, it is possible that some of the activity in the right frontal region may reflect subjects' retrieval of an episodic memory trace associated with recalling at what point in the melody they were supposed to stop. Nonetheless, the major aspect of retrieval elicited by the task should be the semantic component, since the cue sequence only presented the first few notes, and the rest of the tune is stored in long-term memory.

Another brain area active in both of our auditory imagery studies was the SMA. In the current study, the SMA was active in the Cue/Image - Control subtraction (Fig 2, top), and in the subtraction related to generation (Control/Image - Control; Fig 2, bottom), but not in the subtraction related to retrieval. We infer, therefore, that the SMA may be important in the generation of the auditory image. SMA is thought to be important in organization of motor codes, implying a close relationship between auditory and motor memory systems. In our previous paper (Zatorre *et al.*, 1996) we raised the possibility that the activation of SMA may imply a 'singing to oneself' strategy during auditory imagery tasks. Because that study used songs with words, we could not tell if the motor component was related to verbalization or vocalization planning, or both. The tunes in the current study had no lyrics, thus the SMA activation cannot solely reflect preparation of words associated with retrieved tunes. SMA activation seems to reflect motor planning associated with a subvocal singing or humming strategy during the generation process.

We note that SMA activation was found by Rao *et al.* in several of their conditions (Rao *et al.*, 1997). They distinguished between activation of the pre-SMA (positive *y* coordinates) and SMA proper (negative *y* coordinates). The former was active during a pitch discrimination task and the latter was active during the simpler continuation task. In our task, SMA coordinates corresponded to pre-SMA, which Rao *et al.* claim is associated with more complex processing. This conclusion is consistent with the fact that our image-generation task was more complex than their task of imagining an isochronous single tone.

Remaining Issues

One puzzling aspect of our data pertains to the Control/Imagery - Control subtraction. Contrary to our expectations, we did not find activation of the right STG in this condition, which we had assumed would be associated with the phenomenological aspect of seeming to hear an auditory image. Even re-evoking an unfamiliar short sequence of tones just heard ought to require imagery, although perhaps with less vivid imagery than the task of completing a familiar tune given its first few notes. Thus, it is possible that the task may not have elicited a sufficiently lengthy or strong imagery process to be detected by our methods.

A second puzzling aspect of this comparison was the activation in several left frontal sites. The task would appear to require auditory working memory, as the patterns to be imagined were all novel. Previous studies of verbal or figural working memory have implicated dorsolateral regions of the frontal lobe bilaterally in working memory (Petrides *et al.*, 1993; Braver *et al.*, 1997); studies in which tonal working memory was specifically examined have also reported activity within left frontal cortex,

but have generally found much more extensive activity in right frontal sites (Binder *et al.*, 1997; Zatorre *et al.*, 1994). We speculate that some of the areas activated in the left frontal lobe in the present study are related to working memory, and that the SMA is specifically involved in a motor process relevant for auditory image generation, irrespective of the familiarity of the imagined stimulus.

Notes

We gratefully acknowledge the assistance of Dr A.C. Evans and the staff of the McConnell Brain Imaging Center, and of the MNI Cyclotron Unit. We thank Stefan Köhler for helpful comments. This research was supported by grants from the Medical Research Council of Canada (MT11541) and by the McDonnell-Pew Cognitive Neuroscience Program.

Address correspondence to A.R. Halpern, Psychology Department, Bucknell University, Lewisburg, PA 17837, USA. Email: ahalpern@bucknell.edu.

References

Binder J, Frost J, Hammeke T, Cox R, Rao S, Prieto T (1997) Human brain language areas identified by functional magnetic resonance imaging. *J Neurosci* 17:353–362.

Buckner R, Raichle M, Petersen S (1995) Dissociation of human prefrontal cortical areas across different speech production tasks and gender groups. *J Neurophysiol* 74:2163–2173.

Braver TS, Cohen JD, Nystrom LE, Jonides J, Smith EE, Noll DC (1997) A parametric study of prefrontal cortex involvement in human working memory. *NeuroImage* 5:49–62.

Celesia G (1976) Organization of auditory cortical areas in man. *Brain* 99:403–414.

Collins D, Neelin P, Peters T, Evans AC (1994) Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *J Comput Assist Tomogr* 18:192–205.

Démonet JF, Price C., Wise R, Frackowiak RSJ (1994) A PET study of cognitive strategies in normal subjects during language tasks. *Brain* 117:671–682.

Divenyi P, Robinson A (1989) Nonlinguistic auditory capabilities in aphasia. *Brain Lang* 37:290–326.

Evans A, Marrett S, Neelin P, Collins L, Worsley K, Dai W, Milot S, Meyer E, Bub D (1992) Anatomical mapping of functional activation in stereotactic coordinate space. *NeuroImage* 1:43–53.

Farah MJ (1988) Is visual imagery really visual? Overlooked evidence from neuropsychology. *Psychol Rev* 95:307–317.

Farah MJ (1989) Mechanisms of imagery-perception interaction. *J Exp Psychol: Hum Percept Perform* 15:203–211.

Farah MJ (1995) The neural bases of mental imagery. In: *The cognitive neurosciences* (Gazzaniga MS, ed.), pp. 963–975. Cambridge, MA: MIT Press.

Farah MJ, Smith AF (1983) Perceptual interference and facilitation with auditory imagery. *Percept Psychophys* 33:475–478.

Farah MJ, Weisberg LL, Monheit M, Peronnet F (1989) Brain activity underlying mental imagery: event-related potentials during mental image generation. *J Cogn Neurosci* 1:302–316.

Gabrieli JDE, Desmond JE, Domb JB, Wagner AD, Stone MV, Vaidya CJ, Glover GH (1996) Functional magnetic resonance imaging of semantic memory processes in the frontal lobes. *Psychol Sci* 7:278–283.

Galaburda AM, Sanides F (1980) Cytoarchitectonic organization of the human auditory cortex. *J Comp Neurol* 190:597–610.

Goldenberg G, Podreka I, Steiner M, Willmes K, Suess E, Deecke L (1989) Regional cerebral blood flow patterns in visual imagery. *Neuropsychologia* 27:641–664.

Hubbard TL, Stoeckig K (1988) Musical imagery: generation of tones and chords. *J Exp Psychol: Learn Mem Cog* 14:656–667.

Halpern AR (1988) Mental scanning in auditory imagery for tunes. *J Exp Psychol: Learn Mem Cog* 14:434–443.

Janata P (1999) Brain electrical activity evoked by imagined musical events (in press).

Klein D, Milner B, Zatorre RJ, Evans AC, Meyer E (1995) The neural

substrates underlying word generation: a bilingual functional imaging study. *Proc Natl Acad Sci USA* 92:2899–2903.

Kosslyn SM, Alpert NM, Thompson WL, Maljkovic V, Weise SB, Chabris CF, Hamilton SE, Rauch SL, Buonanno FS (1993). Visual mental imagery activates topographically organized visual cortex: PET investigations. *J Cog Neurosci* 5:263–287.

Liégeois-Chauvel C, Peretz I, Babai M, Laguitton V, Chauvel P (1998) Contribution of different cortical areas in the temporal lobes to music processing. *Brain* 121:1853–1867.

Mellet E, Petit L, Mazoyer B, Denis M, Tzourio N (1998) Reopening the mental imagery debate: lessons from functional anatomy. *NeuroImage*, 8:129–139.

Milner BA (1962) Laterality effects in audition. In: *Interhemispheric relations and cerebral dominance* (Mountcastle V, ed.), pp. 177–195. Baltimore, MD: Johns Hopkins Press

Nyberg L, Cabeza R, Tulving E (1996) PET studies of encoding and retrieval. *Psychon Bull Rev* 3:135–148.

Penhune VB, Zatorre RJ, MacDonald JD, Evans AC (1996) Interhemispheric anatomical differences in human primary auditory cortex: probabilistic mapping and volume measurement from magnetic resonance scans. *Cereb Cortex* 6:661–672.

Penhune VB, Zatorre RJ, Evans AC (1998) Cerebellar contributions to motor timing: a PET study of auditory and visual rhythm reproduction. *J Cog Neurosci* 10:752–765.

Petersen S, Fox P, Posner M, Mintun M, Raichle M (1988) Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature* 331:585–589.

Petrides M, Pandya DN (1988) Association fiber pathways to the frontal cortex from the superior temporal region in the Rhesus monkey. *J Comp Neurol* 273:52–66.

Petrides M, Alivisatos B, Meyer E, Evans AC (1993) Functional activation of the human frontal cortex during the performance of verbal working memory tasks. *Proc Natl Acad Sci USA* 90:878–882.

Raichle M, Martin W, Herscovitch P, Mintun M, Markham J (1983) Brain blood flow measured with intravenous O15 H2O. I. Theory and error analysis. *J Nucl Med* 24:790–798.

Rao SM, Harrington DL, Haaland KY, Bobholz JA, Cox RW, Binder JR (1997) Distributed neural systems underlying the timing of movements. *J Neurosci* 17:5528–5535.

Robin DA, Tranel D, Damasio H (1990) Auditory perception of temporal and spectral events in patients with focal left and right cerebral lesions. *Brain Lang* 39:539–555.

Romanski LM, Bates JF, Goldman-Rakic PS (1999) Auditory belt and parabelt projections to the prefrontal cortex in the Rhesus monkey. *J Comp Neurol* 403:141–157.

Samson S, Zatorre RJ (1994) Contribution of the right temporal lobe to musical timbre discrimination. *Neuropsychologia* 32:231–240.

Smith EE, Jonides J (1997) Working memory: a view from neuroimaging. *Cog Psychol* 33:5–42.

Smith EE, Jonides J, Koeppel RA (1996) Dissociating verbal and spatial working memory using PET. *Cereb Cortex* 6:11–20.

Talairach J, Tournoux P (1988) *Co-planar stereotaxic atlas of the human brain*. New York: Thieme.

Westbury C, Zatorre RJ, Evans A (1999) Quantifying variability in the planum temporale: a probability map. *Cereb Cortex* (in press).

Worsley K, Evans A, Marrett S, Neelin P (1992) A three-dimensional statistical analysis for CBF activation studies in human brain. *J Cereb Blood Flow Metab* 12:900–918.

Zatorre RJ (1985) Discrimination and recognition of tonal melodies after unilateral cerebral excisions. *Neuropsychologia* 23:31–41.

Zatorre RJ (1988) Pitch perception of complex tones and human temporal-lobe function. *J Acoust Soc Am* 84:566–572.

Zatorre RJ, Halpern AR (1993) Effect of unilateral temporal-lobe excision on perception and imagery of songs. *Neuropsychologia* 31:221–232.

Zatorre RJ, Samson S (1991) Role of the right temporal neocortex in retention of pitch in auditory short-term memory. *Brain* 114:2403–2417.

Zatorre RJ, Evans AC, Meyer E (1994) Neural mechanisms underlying melodic perception and memory for pitch. *J Neurosci* 14:1908–1919.

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 Cog Neurosci* 8:29–46.