The present functional magnetic resonance imaging study investigates the neural substrates of relative pitch. Musicians and nonmusicians performed 2 same/different discrimination tasks (simple and transposed melody) that differed in whether they required precise encoding and comparison of relative pitch structure, along with 2 control tasks (rhythm and phoneme). The transposed melody task involved a musical transposition of 4 semitones between the target and comparison patterns, requiring listeners to use interval information rather than the absolute value of the individual pitches. Contrasting the transposed melody to the simple melody condition revealed greater activation in the cortex within the intraparietal sulcus (IPS) bilaterally; control tasks did not elicit significant activity in the IPS. Moreover, a whole-brain voxelwise regression analysis of brain oxygenation level-dependent signal showed that activity within the right IPS predicted task performance for both musicians and nonmusicians specifically in the transposed melody condition. Successful performance of the transposed melody task requires encoding and comparison of auditory patterns having different tonal reference points—thus simple tonal memory is not sufficient. Our results point to a role for the IPS in transforming high-level auditory information. We suggest that this area may support a general capacity for transformation and comparison of systematically related stimulus attributes.

Keywords: auditory, fMRI, music, parietal lobes, relative pitch

Introduction

Transposition is a musical manipulation in which a melody is shifted to a new pitch height while exactly maintaining its pitch interval structure. Because melodies are mainly identified by their relative pitches—that is, the interval structure—listeners can easily recognize a transposed melody regardless of musical experience (Attneave and Olson 1971; Dowling and Harwood 1986). In fact, transposition is often thought to be the most frequently used transformation in music (van Egmond and Povel 1996). Music cognition research by Attneave and Olson (1971) demonstrated that pitch interval relationships are encoded and preserved in memory even when a melody is only ever heard in one key. There is also evidence that relative pitch ability may be inborn, as it is seen in infants (Plantinga and Trainor 2005). Musical training can further enhance the ability to make use of relative pitch information, particularly for unfamiliar melodies (Dowling and Harwood 1986).

Transposition is an example of systematic auditory transformation. Although the domain of the manipulation is particular to music, it may be thought of as analogous to other types of stimulus transformations. In particular, because essential relative relationships within the stimulus are preserved (i.e., pitch intervals, formed from the frequency ratios of the tones involved), while absolute information (the pitch values themselves) is not retained, perception of transposed melodies may be considered similar to visuospatial tasks involving coordinate transformation. Object rotation tasks, for example, require that the absolute state (e.g., size or spatial orientation) of stimuli be disregarded to compare their intrinsic structure (Jordan et al. 2001).

In light of these parallels between relative pitch and spatial transformation, it is reasonable to ask whether they are supported by similar neural substrates. The neural correlates of musical transposition and relative pitch have not yet been studied; therefore, we sought to isolate the neural systems involved in relative pitch processing. We designed 2 melodic discrimination tasks that differ mainly in whether they require encoding of relative pitch structure. In the simple melody task, listeners can take advantage of both absolute and relative pitch cues. The transposed melodic task removes the individual pitch values as a cue and therefore requires pitch interval encoding. Subjects performed 2 other auditory discrimination control tasks, which permitted us to distinguish melody-related activation from general effects of memory, effort, and lower level auditory perception.

We recruited musicians and nonmusicians to ensure a wide range of ability, facilitating parametric analyses of task performance and functional activation. Musicians’ formal training may grant them greater ability in discerning inexact melodic transpositions. On the other hand, because relative pitch encoding is fundamental to casual music listening, we expected a diversity of ability and functional activation within the nonmusician population.

We predicted that all tasks would of course evoke activation in auditory cortical areas in the temporal lobes and also in frontal regions related to working memory and other non-specific task demands. But we were most interested in what other neural substrates there might be for these tasks, especially when transposition obliges relative pitch processing because this aspect of melodic cognition has not yet been studied with functional magnetic resonance imaging (fMRI). In particular, we expected that the relative pitch demands of the transposed melody task would recruit additional cortical areas above and beyond those required for simple melody processing, perhaps involving subregions of auditory cortex (Stewart et al. 2008).

Materials and Methods

Subjects

We recruited 2 groups of healthy, right-handed volunteers with normal hearing. A detailed self-reported history of musical training and other
musical experience was obtained from each subject, including estimates of practice hours per week for each year or phase of the participant’s musical activities (as applicable). This information was used to calculate a cumulative measure of hours of musical practice for each subject as described below. Individuals in the musician group (9 subjects, 5 females, mean age 27 years) had a minimum of 8 years of musical training (mean 17) and averaged about 17,300 h of reported lifetime musical practice. Individuals in the nonmusician group (11 subjects, 6 females, mean age 24 years) had no formal musical training and averaged 9 h of lifetime musical experience. All participants gave their informed consent. Ethical approval was granted by the Montreal Neurological Institute Ethics Review Board.

**Tasks**

The subjects performed 4 same-different auditory pattern discrimination tasks: simple melody, transposed melody, rhythm, and phonemes (see example stimuli in Fig. 1). Individual trials consisted of 2 stimulus pattern presentations; subjects judged whether the 2 patterns were the same or different and indicated their response with the left or right button of a computer mouse.

Stimulus durations were varied within each task, and the distribution of durations was matched among conditions. Our study included both musicians and nonmusicians, and we desired a broad range of trial difficulties to provide sensitivity across the full range of musical experience. Varying the number of elements (notes or phoneme sounds) among trials was a straightforward means of ensuring a sufficient range of difficulty. It was important that the tasks be sensitive across the range of ability level present in our sample of musicians and nonmusicians as this would permit us to examine covariation between regional cortical activation and subject task performance. These tasks were also adapted for use in the MRI scanner, as described later.

**Simple Melody**

Stimuli in the simple melody task consisted of unfamiliar melodies in the western major scale, ranging from 5 to 15 notes in duration. The melodies were played with low pass-filtered harmonic tones, using pitches between C4 and E6. All notes were 320 ms in duration, equivalent to eighth notes at a tempo of 93.75 beats per minute. On half the trials, the pitch of a single note anywhere in the melody was changed by up to 5 semitones (median of 2 semitones). The change maintained the key of the melody as well as the melodic contour (the order of upward and downward pitch movement in a melody without regard to magnitude).

**Transposed Melody**

The transposed melody task differed from the simple melody task in 2 ways: 1) all notes of the second stimulus pattern were transposed 4 semitones higher in pitch (in both “same” and “different” trials) and 2) in “different” trials one note was altered by 1 semitone to a pitch outside the pattern’s new key, maintaining the melodic contour. Contour is a particularly salient cue for detecting melodic alterations, and by removing individual pitches and melodic contour as cues, subjects are left only with pitch interval structure as the basis for comparing melodies (Dowling and Harwood 1986). The transposition distance also has an effect on the difficulty of the discrimination. Generally speaking, the more pitches shared between keys the more difficult it is to identify inexact transpositions that mimic the melodic contour (Dowling and Harwood 1986). A transposition of 4 semitones results in 3 tones (out of 7) shared between the scales of the 2 keys. We found this to be a good balance of difficulty for both musicians and nonmusicians.

**Rhythm**

Stimuli in the rhythm task consisted of rhythmic patterns ranging from 5 to 11 notes in duration, played using a single tone at C5 (523 Hz). The patterns sometimes contained one or more rests (pauses between notes) to generate rhythmic variation. The unaltered rhythmic patterns were often syncopated but had underlying metric structure and could be noted by conventional musical notation. The average interval between note onsets was 390 ms. In half the trials (“different” trials), the time offset and/or duration of one or more notes was altered in the second pattern. Subjects were instructed to listen for any difference in timing.

**Phoneme**

The stimuli in the phoneme task were patterns of real speech consonant-vowel syllables (e.g., “a”) spoken in a monotone. Patterns ranged from 5 to 13 elements in length. The interval between phoneme onsets was 320 ms. The full set of phonemes consisted of 12 permutations of 8 consonants [b, k, f, n, p, r, s, j] and 4 vowel sounds [a, u, i, l]. The phonemes were selected to have minimal semantic association. In half the trials, one of the elements in the second pattern was changed to a different phoneme. The 2 stimuli in each trial always used different source recordings (of the same speaker), so that acoustic cues unrelated to phoneme identity could not be used as cues in making the same-different judgment.

**Behavioral Task Procedure**

Tasks were administered by a computer running presentation software (Neurobehavioral Systems, http://www.neurobs.com/) during a single session in a sound-treated room. Stimuli were presented via Sennheiser HD-280Pro headphones driven by a Mackie mixer. Subjects completed 48 6-trial blocks of each task. Task order was counterbalanced, and trials were randomized within each block. This session preceded the fMRI session by a week to 6 months.

**fMRI Parameters**

Subjects underwent functional imaging in a 1.5-T Siemens Sonata with a 1-channel quadrature coil. A 1 × 1 × 1-mm high-resolution T1 anatomical scan was acquired before the 2 functional runs. For fMRI, an echo-planar imaging T2* protocol with a voxel size of 5 × 5 × 5 mm was used to measure brain oxygenation level-dependent (BOLD) signal. We used a sparse sampling (time repetition = 12.5 s) paradigm to minimize any BOLD effect or auditory masking due to MRI scanning noise (Belin et al. 1999; Hall et al. 1999).

The tasks were presented in randomized blocks totaling 24 trials (4 blocks of 6 trials per run) per condition. A subset of trials was selected from each of the tasks as developed for the behavioral testing, omitting the shortest and longest stimuli to accommodate the fixed fMRI scanning interval; trials were selected so that scores on each task would be equivalent to scores obtained in the behavioral session. Melody stimuli ranged in length from 7 to 10 notes, for a total duration of 2.2–3.2 s. The rhythm stimuli ranged from 5 to 10 notes with a total duration of 2.1–3.2 s. Phoneme stimuli ranged from 7 to 10 syllables (2.2–3.1 s). The distribution of pattern durations was matched among all tasks, making the total amount of stimulus energy equivalent in each condition. Twelve silence trials were inserted randomly among the task trials in each run.
A nondiscrimination auditory control task was also included, in which subjects heard 2 equal-length patterns of 320-ms notes at a pitch of C5 and were instructed to click the left button following the second stimulus. The distribution of stimulus durations was matched among all conditions.

A diagram of the stimulus and scan timing is found in Figure 2. The interstimulus interval (ISI) was 1 s for the 2 melody tasks and the phoneme task; in the rhythm task, the ISI was increased to 1.7 s to avoid ambiguity about when the pattern had ended. Because the stimuli varied in duration, a variable delay preceded each stimulus presentation, making the interval between the end of the second stimulus pattern and the beginning of the MRI scan a constant 1050 ms. Our intent was to image brain activity associated with listening, comparing, and—in “different” trials—detecting differences between the 2 stimuli in each trial. The timing of stimuli in our experiment ensured that each fMRI scan was temporally aligned to the end of the comparison pattern such that scan acquisition begins a fixed 1050 ms after the last stimulus.

They received no feedback about their responses.

fMRI Analysis

BOLD signal images were smoothed using a Gaussian kernel (8-mm full width half-maximum) and motion corrected to the second frame of each run using AFNI software (Cox 1996). Then the data were statistically analyzed using a suite of Matlab programs (fmristat, available at www.math.mcgill.ca/keith/fmristat/). Each subject run was fit to a linear model that accounted for stimulus conditions set up in a design matrix corresponding to each acquisition, temporal drift, and temporally correlated errors (Worsley et al. 2002). This yielded the effects variance divided by the fixed-effects variance) in order to stabilize the variance estimate and increase the degrees of freedom. The combined functional runs of each subject were then resampled to a 2 × 2 × 2-mm voxel size and transformed to standard stereotactic space (ICBM-152) (Mazziotta et al. 2001), using a nonlinear function derived from transforming each individual subject’s anatomical scan to a symmetrical version of the ICBM-152 template. In a final step, the standardized, combined runs of all 20 subjects were pooled, yielding combined statistical t maps for each contrast. The threshold for significance was set at $t_{199} = ± 3.94$ (corrected for multiple comparisons $P < 0.05$) for a whole-brain search (240 403 8-mm3 voxels).

Behavioral Results

Subjects were scored on each task based on the percentage of correct responses. During the behavioral session, subjects were required to make a response to proceed to the next trial; in the fMRI session, an omitted response was counted as incorrect. The sensitivity of the musical tasks (simple melody, transposed melody, and rhythm) to musical experience was confirmed during a pilot session with a separate sample of musicians and nonmusicians ($N = 13$). Percent correct performance on each of these tasks correlated significantly with the logarithm of lifetime practice hours ($P < 0.05$). Performance on the phoneme task was not correlated with musical experience.

In the full study of 20 subjects, musicians performed better than nonmusicians on all 3 musical tasks (Fig. 3; 1-tailed Student’s t-test, simple melody, $P < 0.001$; transposed melody, $P < 0.001$; rhythm, $P = 0.001$). Musical training accounted for more than 68% of the variance in performance on both the simple melody and the transposed melody tasks. In the musicians group, the simple melody task was the easiest and the phoneme task was most difficult; for nonmusicians, the phoneme task was the easiest and the transposed melody task the most difficult. As we expected, there was no correlation between musical training and performance on the phoneme task nor did musicians perform significantly better on this task further adding the interaction term $X_iR_iB_j$ to make the complete model $Y_{ij} = X_iB_{ij} + R_iB_{ij} + X_iR_iB_j + e_{ij}$.

Following the individual regression analysis of each run, the 2 runs for each subject were combined using the effects and standard deviations from the previous analysis. This step involved using a mixed-effects model (Gaussian smoothing the ratio of the random-effects variance divided by the fixed-effects variance) in order to stabilize the variance estimate and increase the degrees of freedom. The combined functional runs of each subject were then resampled to a 2 × 2 × 2-mm voxel size and transformed to standard stereotactic space (ICBM-152) (Mazziotta et al. 2001), using a nonlinear function derived from transforming each individual subject’s anatomical scan to a symmetrical version of the ICBM-152 template. In a final step, the standardized, combined runs of all 20 subjects were pooled, yielding combined statistical t maps for each contrast. The threshold for significance was set at $t_{199} = ± 3.94$ (corrected for multiple comparisons $P < 0.05$) for a whole-brain search (240 403 8-mm3 voxels).

![Figure 2. Timeline of events during each fMRI trial. The target auditory pattern (melody, rhythm, or phoneme pattern) is followed by a comparison pattern that is identical or an altered version of the target, and listeners make a same or different judgment. The patterns vary in duration from trial to trial; their presentation is temporally aligned to the end of the comparison pattern such that scan acquisition begins a fixed 1050 ms after the last stimulus.](image)

![Figure 3. Task performance of musicians and nonmusicians in the behavioral and fMRI sessions (SM: simple melody; TM: transposed melody; Rhy: rhythm; Phon: phoneme).](image)
(t-test, $P = 0.35$), indicating that it was an adequate control task for the variable of musical training.

Subjects' task performance in the MRI scanner was very similar to their performance in the behavioral session (comparing between sessions: transposed melody, $r = 0.82$, $P < 0.001$; simple melody, $r = 0.84$, $P < 0.001$; rhythm, $r = 0.76$, $P < 0.001$; phoneme, $r = 0.74$, $P < 0.001$).

**fMRI Results**

**Common Areas of Activation and Deactivation**

We examined all 4 pattern tasks for task-related BOLD activity, using the auditory control condition as baseline (Fig. 4; Supplementary Tables S1–S4). Common areas of task-related activation included bilateral superior temporal lobe, bilateral ventrolateral frontal cortex (VLFC), and other structures that play a role in both motor and cognitive processes in musical tasks (Chen et al. 2008; Leaver et al. 2009): bilateral anterior precentral gyrus, cerebellum, and supplementary motor area. Activity within these regions did not covary with subjects' task performance.

As expected, all 4 tasks evoked significant bilateral BOLD response in temporal auditory areas. In the rhythm task, activation was confined to the planum temporale and was significant only on the left side. In the simple melody, transposed melody, and phoneme tasks, further activation was present in Heschl's gyrus and continued anteriorly along the superior temporal gyrus to around $y = 6$.

The transposed melody and rhythm tasks also evoked significant activation in left primary visual cortex (V1). A similar pattern was seen in the simple melody and phoneme tasks but was just below significance ($t = 3.7$); the magnitude of the BOLD effect in V1 was similar in all 4 tasks. No V1 activation was present in a contrast of the auditory control condition with the silence baseline. No significant recruitment of the intraparietal sulcus (IPS) was observed in contrasts of task versus auditory control, unlike the other contrasts we report below.

Finally, looking at the pattern of BOLD decrease, all 4 tasks evoked bilateral deactivation in areas typically associated with the "default-mode network" (Raichle et al. 2001; Greicius et al. 2003), including the posterior cingulate/precuneus regions, anterior cingulate cortex, inferior parietal cortex bilaterally, and medial prefrontal cortex. Although the precuneus has been reported to be involved in pitch processing as well as a number of other functions (Cavanna and Trimble 2006), here we only find it to be deactivated in a nonspecific fashion.

**BOLD Response Attributable to Relative Pitch**

We followed 2 strategies to find BOLD activity specific to relative pitch task demands. Because the principal difference between the 2 melody tasks is the requirement for relative pitch processing in the transposed melody task, subtracting the simple melody condition from the transposed melody condition reveals activity specific to the difference in relative pitch demands. The result of this contrast analysis is shown in Figure 5, which indicates bilateral elevated activity in the IPS. A stripe of significantly elevated BOLD signal ($P < 0.05$ corrected) lies along the left IPS ($y = –60$ to –38 mm; peak $t = 5.28$ at $x = –40$, $y = –60$, $z = 54$), and there is another focus of elevated activity in the right IPS ($x = 50$, $y = –42$, $z = 56$; $t = 6.25$) (Table 1). When we extract the BOLD signal values at these 2 locations (Fig. 5), we see that the effect on the left IPS can be largely accounted for by a relative decrease within this region in the control task; however, the right IPS showed clear activation not related to any decrease in the control condition. Because performance is greater in the simple melody task than the transposed melody task, we wanted to test whether the BOLD signal increase we see in the IPS can be attributed to a general effect of effort, such as increased attention during the more difficult task. The most difficult condition for participants in the musicians group was the phoneme task; the easiest was the simple melody task. Contrasting these 2 conditions within the musicians group yielded no significant differences in or near the IPS. This finding would suggest that difficulty alone is not driving the IPS activation.

A second approach to identify neural activity specific to relative pitch processing was to ask in which brain areas BOLD signal predicts behavioral performance on the transposed melody task. To this end, we performed a whole-brain
voxel-wise covariation analysis of BOLD signal as a function of percent correct score. Table 2 and Figure 6 show that this analysis yielded 2 foci, 1 bilaterally in anterior IPS (significant peak at \(x = 44, y = -36, z = 40, t = 6.50\); nonsignificant peak of \(t = 3.77\) at the corresponding location in the left hemisphere). This peak is less deep within the sulcus and lies 6 mm more posterior than the right IPS locus found in the contrast analysis; based on its coordinates, the covariation focus appears to lie near or within hIP2 (with 50% probability), a human cytoarchitectonically mapped area believed to form part of the human homolog of the macaque anterior intraparietal area (AIP) (Choi et al. 2006). Another correlation was observed in the right globus pallidus (\(x = 18, y = -4, z = 2; t = 4.98\)); however, this was not matched by any comparable result in the contrast analysis. The significance of this latter result remains to be determined, given that basal ganglia responses in auditory tasks are typically associated with timing-related processes (Peretz and Zatorre 2005).

Subjects’ average transposed melody BOLD signal (relative to auditory control) was extracted at the right IPS voxel, so we could further examine the distribution of BOLD signal versus task performance; the resultant scatter plot is shown in the lower panel of Figure 6. The overall Pearson’s correlation coefficient is \(r = 0.79\), and the correlation between BOLD signal and performance also holds for each group separately (musicians, \(r = 0.44\); nonmusicians, \(r = 0.82\)). In addition to a strong linear relationship between BOLD and performance, the figure shows that equivalently performing musicians and nonmusicians have similar BOLD signal values. The covariation effect in IPS is specific to the transposed melody task; neither IPS voxel’s activity predicts performance in the other conditions (simple melody, rhythm, or phoneme). This specificity is important to note because BOLD signal in the right IPS is also somewhat elevated relative to the control condition in both the rhythm and the phoneme conditions (Fig. 5). However, there is no covariation with task performance in these conditions.

For this analysis, we used the behavioral scores collected outside the magnet in a separate test session. Correlations between performance and BOLD signal were stronger with these behavioral session data than when the in-scanner data were used (e.g., in the transposed melody task, the peak correlation coefficient was 0.79 vs. 0.66). This is probably due to

Table 1

<table>
<thead>
<tr>
<th>Region</th>
<th>xyz</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right IPS</td>
<td>50</td>
<td>-42</td>
</tr>
<tr>
<td>Left IPS</td>
<td>-40</td>
<td>-60</td>
</tr>
<tr>
<td>Left anterior cingulate gyrus</td>
<td>-6</td>
<td>28</td>
</tr>
</tbody>
</table>

Note: A threshold of \(P < 0.05\) (corrected) was applied; when local peaks were closer than 10 mm, only the most significant peak is reported.

Table 2

<table>
<thead>
<tr>
<th>Region</th>
<th>xyz</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right IPS</td>
<td>44</td>
<td>-36</td>
</tr>
<tr>
<td>Right globus pallidus</td>
<td>18</td>
<td>-4</td>
</tr>
</tbody>
</table>

Note: A threshold of \(P < 0.05\) (corrected) was applied; when local peaks were closer than 10 mm, only the most significant peak is reported.
nonmusicians performed well below the musicians, whereas others had performance comparable to lower performing musicians. Likewise, musical experience and task performance varied within the musician group. To use this variation to our advantage, we performed a covariance analysis of BOLD signal and musical experience (logarithm of lifetime hours of experience) to determine if the IPS result identified above was related to training. Notably, there was no covariation between BOLD signal in the parietal lobes and musical experience during the transposed melody task or on any of the other 3 tasks. This analysis did reveal significant positive interactions between musical experience and BOLD signal in the transposed melody task in the right precentral gyrus ($x = 28, y = -10, z = 50; t = 5.08$). There was also a near-significant effect in the simple melody task (right precentral gyrus at $x = 28, y = -14, z = 48; t = 4.86$). No experience-BOLD signal interaction effects were found in the rhythm or phoneme tasks.

**Functional Connectivity with Right IPS**

Because the IPS has anatomical connectivity with auditory and visual regions (Lewis and Van Essen 2000; Schroeder and Foxe 2002; Frey et al. 2008), we used functional connectivity analyses to test whether activation in the right IPS is temporally correlated with BOLD signal in these areas. Two seed voxels were chosen in right IPS: one at $x = 44, y = -36, z = 40$, corresponding to the right hemisphere IPS peak in the transposed melody task covariation result, and another at $x = 50, y = -42, z = 56$ corresponding to the transposed versus simple melody task contrast. In the baseline functional connectivity analysis, which finds functional connectivity with the seed voxel independent of the task condition, both seed voxels’re activity positively correlated with activity in the contralateral IPS, bilateral precentral gyrus, bilateral VLFC, and right dorsolateral frontal cortex (DLFC); there were also negative correlations in ventromedial frontal cortex, posterior cingulate cortex, and inferotemporal areas (Supplementary Tables S5–S6). A second functional connectivity analysis tested where functional connectivity is modulated by the task condition, comparing simple to transposed melody. In the task-modulated analysis with the 44, -36, 40 seed voxel, there was a significant voxel in left hippocampus ($x = -20, y = -16, z = -24; t = 5.15$) and a nonsignificant peak in the left basal ganglia. A task-modulated analysis using the $x = 50, y = -42, z = 56$ seed voxel found no significantly correlated activity and nonsignificant peaks in right IPS and medial parietal areas. None of the functional connectivity analyses found a significant correlation in occipital visual areas.

**Discussion**

In this study, we investigated the neural responses associated with musical transposition, an operation that requires relative pitch processing and constitutes a type of systematic auditory transformation. Compared with conditions that did not require relative pitch processing, the transposed melody task evoked elevated and performance-dependent activity in the IPS. The present findings therefore raise questions about the role that posterior parietal areas may play in manipulating auditory information. Certain stimulus attributes and cognitive demands were designed to overlap among the 4 task conditions, so that by examining the imaging results in relation to these task parameters we can gain insight into the specific role of the IPS in relative pitch processing.
Task-Related Activation

Our melody and control tasks have in common several important processing demands, including auditory perception, working memory, a discrimination judgment, and response selection. These cognitive operations were evident in activation patterns that were similar across conditions, particularly in auditory cortex and VLFC (Fig. 4). All 4 tasks activated either primary auditory cortex or the planum temporale. Activation of auditory areas was considerably more extensive for the melody and phoneme tasks than the rhythm task, which only activated the planum temporale. This is consistent with literature reporting greater anterior and posterior temporal activation for auditory stimuli having melodic or timbral structure (e.g., Patterson et al. 2002).

There is significant VLFC activation in all tasks, including the phoneme task. We attribute this activity to active retrieval from working memory (Petrides 2002; Kostopoulos and Petrides 2003), which would be required by the nature of the comparison task. Similar VLFC recruitment has been noted in other same-different melodic comparison tasks (Zatorre et al. 1994; Griffiths et al. 1999).

Anatomical connectivity of the IPS (Schmahmann et al. 2007) would lead us to predict that parietal activity during working memory tasks be accompanied by activation of DLFC; these 2 regions often act in concert during working memory tasks (Champod and Petrides 2007). Although we did not observe DLFC activation in contrasts of task versus auditory control, an analysis of baseline functional connectivity showed that BOLD signal in VLFC and right IPS is temporally correlated. However, this functional connectivity was not significantly modulated by task condition, indicating a similar interaction between these regions across tasks.

The Role of the IPS in Melodic Transposition

The IPS is an area long associated with higher order visual processing, particularly for spatial operations (Husain and Nachev 2007); it is also implicated in arithmetic and quantity processing (Kong et al. 2005; Ischebeck et al. 2006; Piazza et al. 2007) Activation in the IPS is frequently reported in auditory paradigms, both for spatial operations (Weeks et al. 2000; Alain et al. 2001; Maeder et al. 2001; Zatorre, Bouffard, et al. 2002) and also in nonspatial tasks (Gaab et al. 2003; Cusack 2005; Zarate and Zatorre 2008). Although we had not specifically predicted IPS involvement in the transposed melody task, the contrast analysis indicated a significant effect in this region (Fig. 5). Perhaps, more importantly, the whole-brain performance versus BOLD signal covariance analysis (Fig. 6) revealed a highly significant effect in right IPS for the transposed melody task but not for the other tasks. It is important to note that the latter analysis was unconstrained with regard to anatomical location because it was performed on a whole-brain basis and not on a predetermined region of interest. The fact that both analyses converge on the IPS, particularly on the right side, therefore constitutes good evidence of the specific recruitment of this region. The right-sided predominance in both analyses would be consistent with a great deal of evidence for specialization of right hemisphere mechanisms in melody processing (Zatorre, Belin, and Penhune 2002).

The transposition task is of particular interest because of its reliance on precise relative pitch ability, so finding a task-specific effect in the IPS was thought provoking. Several hypothetical lines of explanation come to mind: 1) the IPS recruitment is related to visual imagery; 2) the parietal lobe activity is related to attentional demands or effortfulness of the task; 3) the parietal lobe activity is related to working memory; and 4) the IPS handles abstract operations like transformations of frames of reference. We favor the fourth interpretation but will consider each of these possibilities in turn.

Visual Imagery

IPS recruitment could be due to visual mental imagery (Trotjano et al. 2000). Some types of visual imagery, for example, of a musical score or keys on an instrument, would presumably depend on musical training; hence, activation resulting from such visualization should be present in musicians but not nonmusicians. However, we did not find any effect of musical training on BOLD signal in the IPS either in a group contrast or in a covariation analysis. Moreover, as seen in the BOLD scatter plot (Fig. 6), better performing nonmusicians have greater recruitment of right IPS, the same as musicians. This supports the conclusion that IPS activation in the transposed melody task is not solely a consequence of musical training, and it is therefore not related to strategies that could only follow from musical training, such as visualizing a musical score. Other forms of visual imagery might not rely on musical training. In particular, there is evidence that nonmusicians may make spatial associations with “pitch height” (Rusconi et al. 2005); however, the effect of this simple factor, if present, should be equivalent in our 2 melody tasks. Successfully applying visuospatial strategies to transposition would require a degree of abstraction and precision beyond that demonstrated by Rusconi et al. because pitch height in and of itself is insufficient to perform the transposed melody task, which requires transforming the auditory information into intervals (frequency ratios), independent of specific pitch values.

We did observe recruitment of V1 in all the tasks, which might be related to visual mental imagery (Kosslyn et al. 1999; Ganis et al. 2004). However, unlike IPS activity, V1 activity did not predict performance on the transposed melody task. Furthermore, occipital activation was not unique to any particular task (in fact, it was quite consistent across the tasks). We also failed to find functional connectivity between seed voxels in the IPS and occipital cortex. We therefore consider that whereas the IPS appears to have a special role in melodic transposition, V1 recruitment serves a more general capacity in all 4 tasks; this argues against a role for visual imagery specifically in the process of transposition.

Attention

At first glance, differences in attentional demands could explain the differential activation we see in IPS between the transposed and simple melody tasks because the latter task is easier for both musicians and nonmusicians. However, when we tested for an effect of task difficulty in the IPS by contrasting the tasks with greatest and lowest performance, no difference in BOLD signal was found in or near the IPS nor did number of notes affect BOLD values. Moreover, if greater IPS recruitment in the transposed melody task were indeed driven by attention or effort-related factors, then we would expect higher performing subjects to have lower BOLD signal in the IPS. In fact, we observe the opposite: greater BOLD signal in the IPS is associated with success on the task (Fig. 6). Therefore, although we cannot completely discount the role of attention,
we find little support for the idea that attention or difficulty drives the differential IPS activation we observe in the transposed melody task.

**Working Memory**

With regard to working memory, the neuroimaging literature (e.g., Champod and Petrides 2007; Berryhill and Olson 2008; Wendelken et al. 2008) along with many other sources of data indicates an important role for the posterior parietal cortex in working memory. As the duration and number of elements were closely matched in all 4 tasks, the quantity of items to maintain and monitor should be equivalent; the differences among the tasks lie mainly in the manner in which the auditory information must be manipulated. In a voxel-of-interest analysis, we explicitly verified that BOLD signal was not related to the number of items (notes) in the melody tasks, indicating that working memory load per se did not affect IPS level of activity. Although the number of items was equivalent across the tasks, the transposed melody task always requires tracking of intervals in the 2 compared melodies, whereas in the simple melody task this would not necessarily be required. Therefore, although we cannot completely equate the tasks in terms of working memory, we interpret differences in BOLD signal between tasks as mainly reflecting operations being performed on memory contents (e.g., transposing melodies or transforming them to a relative pitch representation) rather than load due to quantity per se. Champod and Petrides (2007) directly contrasted the manipulation versus monitoring components of working memory using abstract visual stimuli, and they demonstrated that the IPS is more involved when working memory contents must be manipulated. We believe that it is the manipulation of melodies into a relative pitch representation that causes greater IPS recruitment and the correlation between IPS activity and success on the task.

**Transformation**

The human IPS is known to be activated during spatial tasks like visual mental rotation (Jordan et al. 2001; Zacks 2008) and visually guided reaching and grasping (Frey et al. 2005). Studies have shown that the primate parietal sulcus is organized into subregions, many specialized for spatial process that integrate visual, tactile, auditory, and/or motor processing. These include the AIP, posterior reach region, lateral interparietal area (LIP), caudal intraparietal area, and ventral intraparietal area (VIP) (Grefkes and Fink 2005). These areas form part of the "dorsal stream," subserving spatial localization, representation of the external world in body-, head-, or retina-centered space, and control of action (Culham and Kanwisher 2001), and they have been characterized as forming a critical part of a vision for action pathway (Milner and Goodale 2008). It has been further suggested that such parietal areas integrate multimodal feedforward and feedback information for further computation supporting perception (Rauschecker and Scott 2009). Common to these various functions is the recoding of one representation—for example, a retinotopic image of the object to be grasped—into another frame of reference to permit action (e.g., movement of the arm relative to its current location) or comparison. Object rotation tasks, for example, require that the absolute state (size or spatial orientation) of stimuli be disregarded in order to compare their intrinsic structure. In humans, the organization of posterior parietal cortex is less well understood than for nonhuman primates; however, probable homologs of AIP, VIP, and LIP have been proposed in the anterior, ventral, and posterior human IPS, respectively (Culham and Kanwisher 2001; Choi et al. 2006). Human posterior parietal cortex also seems to be involved in nonspatial operations to a greater extent than in the macaque, such as manipulating working memory contents and maintaining and controlling attention (Husain and Nachev 2007).

The human IPS is situated in the middle of a functional gradient spanning from mainly spatial functions in the superior parietal lobe to mainly nonspatial functions in the inferior parietal lobe (Husain and Nachev 2007). It is a "classic" multisensory region, receiving converging anatomical inputs from visual, auditory, and tactile sensory cortices (Schroeder and Foxe 2002; Frey et al. 2008). Our behavioral covariation result appears to lie within or very close to hIP2, a cytoarchitectonically defined region believed to form part of the human equivalent of macaque AIP (Choi et al. 2006). In both humans and the macaque, this area is multimodal and is involved in orientation judgments as well as cross-modal information transfer where frames of reference differ (Grefkes and Fink 2005). The IPS is therefore in a prime position to apply systematic transformations to auditory as well as visual information.

Converging evidence for IPS involvement in musical transformation comes from a recent fMRI study in our laboratory. Zatorre et al. (2009) tested subjects discriminating exact from inexact temporal reversals of melodies. The task explicitly required that subjects imagine what the reversed melody would sound like. Similar to the transposed melody task, this mental melodic reversal produced activation in anterior portions of the IPS, overlapping with those found in the present study. This convergence allows us to ask which task demands may be similar between the experiments. Both tasks involve the transformation of a stimulus representation while maintaining its internal structure; in the melody reversal task, the result is reflected in the time dimension, whereas in the transposed melody task the melody is shifted in the pitch dimension.

Additional converging evidence comes from behavioral studies. In an experiment by Cupchik et al. (2001), subjects performed tasks of mental rotation in 3 dimensions as well as melodic pitch inversion and temporal reversal; subjects' accuracy on the spatial task predicted their performance on the musical judgments, indicating a close functional relationship. This result supports the conclusion that visuospatial transformations and melodic manipulations are linked. Conversely, a recent paper on persons with amusia (Douglas and Bilkey 2007) showed an association between impaired musical ability and impaired visual mental rotation.

These results underline the multimodal and complex nature of parietal function. In particular, although the parietal lobes' capacity to systematically transform stimulus representations is typically understood in the context of visual information, it is evident that the capability extends to other types of stimulus representations—perhaps any stimulus representation depending upon precise relationships among its elements (e.g., distance, pitch interval, or numeric relationship).

**Variability in Relative Pitch Ability in Nonmusicians**

Although there was a positive effect of musical training on task performance (Fig. 3), nonmusicians had a range of performance that overlapped with musicians on the musical tasks. Do the
nonmusicians who perform relatively better have similar patterns of activation as equivalently performing musicians? One answer to that question is found in Figure 6: the performance versus BOLD relationship in the right IPS forms a continuous linear distribution for musicians and nonmusicians. Although musical training is associated with greater performance and greater IPS recruitment, a clear relationship between IPS recruitment and performance is also seen in nonmusicians alone.

It may be noted from Figure 6 that the lowest performing nonmusicians were effectively at chance on the task. However, the BOLD signal in right IPS does not simply distinguish those who can perform the task from those who cannot; rather, taking the 2 groups of participants separately, there is a significant positive relationship between BOLD activation and task performance in each group. Moreover, where musicians and nonmusicians overlap in performance, they also overlap in BOLD signal level, without a discontinuity in the distribution of performance versus BOLD signal.

The wide range of performance among nonmusicians may arise from differences in musical exposure; alternatively, or additionally, these differences may be driven by preexisting factors. The role of "native ability" in skilled pursuits like music is of keen interest in the field of neuroscience. The performance overlap between musicians and nonmusicians indicates that, although important, musical training may not be the sole determinant of relative pitch performance.

Conclusion

The present experiment demonstrates that the cortex within the IPS is recruited when auditory information must be systematically transformed. The transposed melody task taps into a fundamental element of music perception—relative pitch encoding—that is present in untrained individuals and is further refined with musical training. We found that activity in an anterior area of the intraparietal cortex specifically predicts relative pitch ability in musicians and nonmusicians, providing a direct link between neural activity and the function of interest. Our results suggest that the transformational operations resident in the IPS, although commonly applied to visual information, are available to apply to auditory representations.

Supplementary Material

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

Funding

Canadian Institutes of Health Research (11541, 14995); and the Natural Sciences and Engineering Research Council of Canada (217297).

Notes

We thank Patrick Bermudez, Joyce Chen, Marc Schoenwiesner, and the staff of the McConnell Brain Imaging Centre. Conflict of Interest: None declared.

Address correspondence to Nicholas E. V. Foster, Montreal Neurological Institute, Room 276, 3801 University Street, Montreal, Québec H3A 2B4, Canada. Email: foster+17@bic.mni.mcgill.ca.

References


