The Neurocognitive Components of Pitch Processing: Insights from Absolute Pitch

Sarah J. Wilson[1], Dean Lusher[1], Catherine Y. Wan[1], Paul Dudgeon[1] and David C. Reutens[2]

[1]School of Behavioural Science, University of Melbourne, Victoria, 3010, Australia and [2]Southern Health Neurosciences, Monash Medical Centre, Clayton, Victoria, 3168, Australia

The natural variability of pitch naming ability in the population (known as absolute pitch or AP) provides an ideal method for investigating individual differences in pitch processing and auditory knowledge formation and representation. We have demonstrated the involvement of different cognitive processes in AP ability that reflects varying skill expertise in the presence of similar early age of onset of music tuition. These processes were related to different regions of brain activity, including those involved in pitch working memory (right prefrontal cortex) and the long-term representation of pitch (superior temporal gyrus). They reflected expertise through the use of context dependent pitch cues and the level of automaticity of pitch naming. They impart functional significance to structural asymmetry differences in the planum temporale of musicians and establish a neurobiological basis for an AP template. More generally, they indicate variability of knowledge representation in the presence of environmental fostering of early cognitive development that translates to differences in cognitive ability.

Keywords: functional neuroimaging, musical skill expertise, pitch identification, planum temporale, relative pitch perception, structural neuroimaging

Introduction

The skill of perfect or absolute pitch (AP) is typically defined as the ability to identify a pitch (e.g., the musical note C) without using a reference tone. It is commonly associated with the commencement of early musical training during a sensitive period of development, with genetic influences also playing a significant role (Zatorre 2003). It is an important ability to study because it can inform us about interactions between genes and the environment, and specifically how tuition during a sensitive period may lead to changes in brain structure and function that underpin the development of knowledge representations and behavioral skills.

The acquisition of AP may be considered a deviation from the typical development of relative pitch (RP) ability that is used by the majority of individuals to process the relations between tones (Trehub 2003; Trehub and Hannon 2006). Previous research has suggested that AP is an inherited autosomal dominant trait, with strong support for familial aggregation of AP after controlling for early musical training (Profita and Bidder 1988; Baharloo et al. 2000). The penetrance of AP may be influenced by the presence and type of early musical training, implying that in genetically predisposed individuals, the brain may be "particularly amenable to the establishment of new circuits or to the fine-tuning of preexisting circuits involved in pitch perception" (Baharloo et al. 1998, 2000; Gregersen et al. 2001). Although reasons for undertaking musical training may in part reflect genetic predisposition, the impact of the environment and in particular, the role of a sensitive period in the development of neural changes has not yet been fully determined.

A 2-component model of AP ability has been proposed by Levitin (1994; Levitin and Rogers 2005). The first component entails absolute representation of pitch in long-term memory that appears common to all humans and some other species (Terhardt and Ward 1982; Halpern 1989; Saffran and Griepentrog 2001; Deutsch 2002). The second component involves pitch labeling, typically using a verbal code that may be acquired from training during the sensitive period in select individuals. This component has been proposed to involve conditional associative memory, with tones organized in nominal categories with up to 70 categories identified (Zatorre 2003). The model is hierarchical in that the first component constitutes a prerequisite for the second. In AP possessors, AP representation may be integrated with the verbal code, creating a unique pitch template (Levitin and Rogers 2005).

Research has begun to address the neurobiological basis of this model. Functional neuroimaging findings reported by Zatorre et al. (1998) pointed to the importance of left dorsolateral frontal cortex in supporting conditional verbal-pitch associations in AP possessors while listening to tones (an implicit pitch naming task). This task also engaged frontotemporal auditory working memory processes (pitch and verbal) presumably to support the retention of tones "in mind" while associating their verbal labels. Because left dorsolateral frontal cortex has been activated in nonmusicians during a conditional pitch association task, one interpretation is that there is no difference in the brains of musicians with or without AP (Bermudez and Zatorre 2005).

Structural neuroimaging research by Schlaug et al. (1995) has implicated a region in the posterior superior temporal gyrus, the planum temporale, based on greater left-right asymmetry in AP musicians compared with musicians without AP and normal controls (Keenan et al. 2001). The degree of activation of the left planum temporale has been correlated with the age of onset of musical training (Ohnishi et al. 2001), and it has been proposed to be important in matching spectrotemporal patterns with stored patterns or templates of previously experienced sound objects (Griffiths and Warren 2002). In particular, Griffiths and Warren (2002) have likened the planum temporale to a "computational hub," allowing abstraction of properties of complex sounds for further analysis by higher cortical regions (Griffiths and Warren 2002). Taken together, these findings point to a possible role of the planum temporale in the long-term representation of pitch ("AP template"). In light of Zatorre's research, however, there is

a need to directly assess the functional neurobiological correlates of this notion. To achieve this in the current study we investigated the sights of activation along the anterior–posterior extent of the superior temporal gyrus during an explicit pitch naming task.

A key conceptual challenge to studying individuals with unique cognitive abilities is the identification of an appropriate reference group. Examining a group who lack the skill poses difficulties for interpretation because these individuals are presumably unable to perform the task. Traditionally, AP has been contrasted with RP perception, however more recent research has suggested that these 2 skills lie on a continuum with "quasi-absolute pitch" (QAP) falling in between (Levitin and Rogers 2005). Because QAP shows variability in the behavioral phenotype and its level of dependence on the auditory context for pitch cues, it provides a crucial but underutilized paradigm for addressing this issue.

Our first aim was to assess functional involvement of the planum temporale in AP ability. We developed an explicit in-scanner pitch naming task using measures of response accuracy and reaction time after subtracting conditional association and auditory working memory processes. We hypothesized that the long-term representation of AP engages an AP template as evident from activation of the posterior superior temporal gyrus and structural asymmetry of the planum temporale. We further postulated that variability in this template would account for behavioral variability in AP ability. Thus, our second question examined whether functional activation and structural measures of the planum temporale differed for the AP and QAP groups. We predicted that QAP musicians would engage additional neurocognitive processes to assist their more limited AP template, as reflected by their lesser AP skill.

Although RP is common to all individuals, previous research has suggested that the automatic nature of AP may interfere with RP judgments in AP possessors (Miyazaki and Rakowski 2002). The neurocognitive basis of any interference afforded by dual representations has not been investigated (Zatorre 2003). Thus, our third question examined patterns of cerebral activity associated with RP ability in the same musicians with varying AP, as well as a group of musicians with RP. We also assessed the presence of structural differences in the planum temporale of these RP musicians.

Materials and Methods

Subjects

One hundred and twenty-five musicians were recruited from the community and tertiary education settings, and underwent neurological and psychiatric screening, and audiometry for the presence of significant hearing loss. Of these musicians, 12 (6 males) were consecutively assigned to each of the AP, QAP, and RP groups, taking age, sex, handedness, medical history, and willingness to undergo positron emission tomography (PET) into account. This resulted in a sample of 36 musicians with similar demographic and musical background characteristics (see Table 1). The study was approved by the Human Research Ethics Committees of Austin Health and the University of Melbourne, Australia, and all participants gave written informed consent.

Out-of-Scanner AP Task

The out-of-scanner AP task was based on an established behavioral method for identifying AP possessors (Takeuchi and Hulse 1993). The musicians were required to identify a series of 50 randomized piano

Table 1
Characteristics of the musicians

Characteristic	AP ($n = 12$)	QAP ($n = 12$)	RP ($n = 12$)
Sex	6F, 6M	6F, 6M	6F, 6M
Mean age (years $\pm$ SD)	22.8 (4.9)	26.2 (13.8)	26.0 (6.9)
Mean education (years $\pm$ SD)	16.1 (3.1)	15.6 (2.4)	16.8 (2.5)
Mean years playing ($\pm$SD)	15.6 (4.5)	16.8 (14.9)	15.6 (8.2)
First instrument			
Keyboard	10	8	11
Violin	2	2	1
Recorder		2	
Principal instrument			
Keyboard	8	7	6
Strings	2	2	4
Woodwind		1	
Brass	1	1	1
Percussion			1
Voice	1	1	

Note: All musicians were right-handed and were able to read and write music in the treble and bass clefs.

tones by their musical note names (pitch chroma) without feedback of the accuracy of their responses. The tones were selected from the equal tempered scale, ranging from C_2 to C_5 (concert pitch A_4 440 Hz, American notation) and proportionally distributed across the black and white notes of the piano. Each tone had a duration of 500 ms, followed by an interval of 2.5-s response time (1 stimulus = 3 s). The piano tones were synthesized using a "stereo grand" piano timbre of a Yamaha S80, 88-note fully weighted keyboard, linked to a G4 Apple Macintosh computer (Mac OS 8.6). Pro Tools LE software (5.0, Digidesign, Daly City, CA) was used to assemble the stimuli, which were binaurally presented to the musicians via loud speakers at a comfortable listening level in an auditory laboratory. For both the out-of-scanner and in-scanner AP tasks, a piano timbre was used to capture variability in AP ability. In accordance with previous research (Takeuchi and Hulse 1993), musicians were classified as having AP if $\geq 90\%$ of the tones were correctly identified, whereas those identifying $\leq 20\%$ were considered to lack AP ability (RP group). Those falling in between were classified as QAP musicians. Correct pitch identification did not require accurate octave classification and semitone errors were coded as incorrect for all participants.

Cognitive Activation Paradigm

Three types of stimuli were created: noise bursts, piano tones, and piano chords. These were binaurally presented to the musicians during scanning via Ear Tone 3A earplugs at a comfortable listening level. A microphone was used to communicate with the musicians, including delivery of task instructions. The stimuli were played via a Pro 2 stereo audio mixer linked to a PC running stimulus presentation software written in Matlab (5.0, MathWorks, Novi, MI). In each task, a total of 54 stimulus pairs were presented (18 per scan) with an interstimulus interval of 5 s. Participant responses were recorded on a Sony Digital Audio Tape-corder machine (TCD-D10 PRO, Sony, San Diego, CA) to measure response accuracy and reaction time.

For the baseline task, pairs of noise bursts comprised 250 ms of silence, 1 s of noise burst, 250 ms of silence, and 500 ms of noise burst. Noise bursts consisted of white noise with similar onset/offset characteristics to the piano tones and chords, and were created using the "wvnoise" timbre of the S80 keyboard. The baseline task was designed to subtract activation associated with basic auditory and linguistic processing, including conditional association and auditory working memory processes at relatively low processing load. During scanning, the musicians were required to listen carefully to each pair of noise bursts and respond with the words "C natural."

Piano chord and tone pairs for the 2 activation tasks comprised 250 ms of silence, 1 s of arpeggiated piano chord, 250 ms of silence, and 500 ms of piano tone. They were created using the "stereo grand" piano timbre of the S80 keyboard, and were randomly distributed across the equal tempered scale (C_2–C_5). For the in-scanner AP task ("pitch naming") the musicians were presented with an arpeggiated chord of octaves followed by a tone of the same pitch (the target tone) and

required to identify this tone using its full musical note name (e.g., "C natural"; Fig. 1a). If they did not recognize the tone, they said "pass" and moved on to the next item. For each tone, correct octave classification was not required and semitone errors were coded as incorrect.
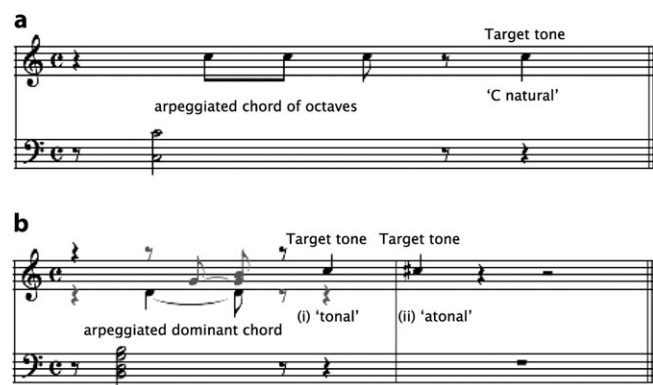
For the RP task ("tonal classification"), musicians were presented with an arpeggiated dominant chord followed by a tone that was either 1) the tonic of the key, or 2) a semitone higher than the tonic (Fig. 1b). Piano chords were distributed across each major key within the range of G#$_1$–A$_4$ of the equal tempered scale, with random modulation of items within this range promoting RP processing. During training, the participants were instructed to identify condition 1) as "tonal" and condition 2) as "atonal." If they were unsure, they said "pass" and moved on to the next item.

The tonal classification task was derived from the well-established probe-tone technique, that is paradigmatic of RP processing tasks in the tonal idiom (Krumhansl 1991). Although AP possessors may perform RP tasks using pitch labeling, the tonal classification task favored a RP processing strategy by promoting a global tonality judgment. In particular, it required the listener to perform an interval judgment between the second last note and the last note of the item within the tonal context established by the preceding intervals of the arpeggiated sequence (Fig. 1b). An interval of one semitone (leading note to tonic) created a sense of tonal resolution ("tonal" condition), whereas the interval of a whole tone left the sequence unresolved ("atonal" condition). In other words, the tonal classification task was purposely designed to be similar to the pitch naming task with the exception that the pitch naming task required an AP judgment labeled according to note name, whereas the tonal classification task favored a RP judgment labeled according to the degree to which the sequence conformed with the Western tonal idiom (Fig. 1). Similar to the baseline task, the tonal classification task also allowed subtraction of basic auditory and linguistic processing, conditional verbal-pitch associations, and auditory working memory functions.

Initial analysis of the behavioral data revealed that they were suitable for parametric analyses that were performed using SPSS 11 (SPSS, Chicago, IL), with $P < 0.05$ (2-tailed) set as the criterion of statistical significance. Independent samples t-tests and analysis of variance (ANOVA) with planned contrasts were used for 2 group or 3 group comparisons respectively. Variances were not assumed for independent samples t-tests when Levene's test of equality of variances was not met. Within group comparisons were performed using paired sample t-tests and repeated measures ANOVA.

Neuroimaging Acquisition and Analysis
PET scans were acquired on an ECAT positron tomograph (951/31R, CTI Siemens, Knoxville, TN). Each musician underwent 9 PET scans following the bolus injection of the blood flow tracer [^{15}O]H$_2$O with an interval of at least 10 min between scans. The total effective dose equivalent of radioactivity for the entire study was less than 5.0 mSv for each participant. Integrated radioactivity counts accumulated over a 60-s acquisition period beginning with the rising phase of head counts was used as an index of perfusion. Three scans were performed for the baseline task and 3 for each of the activation tasks. Performance commenced 30 s before the start of scan acquisition and continued throughout the entire scanning period. Prior to scanning the musicians received training on each task to ensure that they fully understood the task requirements. High-resolution T_1-weighted spoiled gradient echo magnetic resonance (MR) scans (time repetition 35 ms, time echo 7 ms, flip angle 35°, field of view 24 cm) comprising 120 contiguous slices of 1-mm thickness and 1 mm × 1 mm pixel dimension were also acquired in each musician.

For each PET study, subsequent PET images were aligned to the first to correct for head movement between scans using automated image registration software. Both MR and PET images were then stereotactically normalized using an automated algorithm at progressively higher resolutions (Talairach and Tourneaux 1988). Images were blurred with a 16-mm full width at half maximum Gaussian filter to correct for anatomical differences between participants. PET activation data were analyzed using the general linear model (SPM2, Wellcome Department of Imaging Neuroscience, University College London, UK), and a statistical parametric map of the t-statistic for the regression coefficient was generated using contrasts to investigate specific effects. These contrasts were guided by the findings from the behavioral data collected during scanning. Statistical significance was ascertained using distributional approximations based on the theory of Gaussian random fields (Worsley et al. 1992). The threshold for significance was $P < 0.05$ corrected for analysis across the whole brain, unless otherwise stated. For the pitch naming task, region of interest (ROI) analyses were conducted using regions automatically generated from WFU PickAtlas version 2.0 with no dilation (Maldjian et al. 2003). This software allows the generation of ROI masks on the Talairach database (Talairach and Tourneaux 1988). Labels for activation peaks were obtained in Talairach co-ordinates using the Talairach Daemon software, which provides accuracy similar to that of neuroanatomical experts (Lancaster et al. 2000). In addition, the gyral location of the activation peaks was identified on the spatially normalized MR scans of individual participants.

Volumetric Analysis of the Planum Temporale
MR images were registered into standard orientation using an existing voxel-based method (Woods et al. 1993) and anisotropic scaling was performed along the 3 principal axes. The boundaries used to define the planum temporale were validated against human cadaver brains by Steinmetz et al. (1989). The gray matter underlying the planum temporale was manually segmented using interactive mouse driven software (Display, Montreal Neurological Institute, Canada) that enabled simultaneous analysis of coronal, sagittal and axial slices. Volume measurements were based on a voxel counting algorithm that multiplies the sum of the number of voxels contained in the painted volume by the voxel dimensions (0.5 × 0.5 x 1 mm). All analyses were performed blind to the musicians' pitch naming ability. Intrarater reliability for 25% of the scans was measured more than 6 months apart, revealing adequate intrarater agreement (left planum temporale, Spearman's rho = 0.75; right planum temporale, Spearman's rho = 0.63).

Figure 1. Cognitive activation tasks. (a) The pitch naming task showing an arpeggiated chord of octaves followed by the target tone of the same pitch that required an AP judgment using the full musical note name. (b) The tonal classification task showing an arpeggiated dominant chord followed by the target tone that favored a RP judgment reflecting the degree to which the target tone completed the musical phrase ("tonal" vs. "atonal").

Results

Pitch Naming
Clustered box plots of the accuracy of pitch naming revealed consonance between a priori categorization of musicians into AP, QAP, and RP groups and in-scanner AP performance (Fig. 2a), with a high correlation between our in-scanner and out-of-scanner AP tasks (Pearson $r = 0.963$, $P < 0.001$). In light of the well-established nature of the out-of-scanner task (Takeuchi and Hulse 1993), these findings support the validity of our in-scanner AP paradigm.
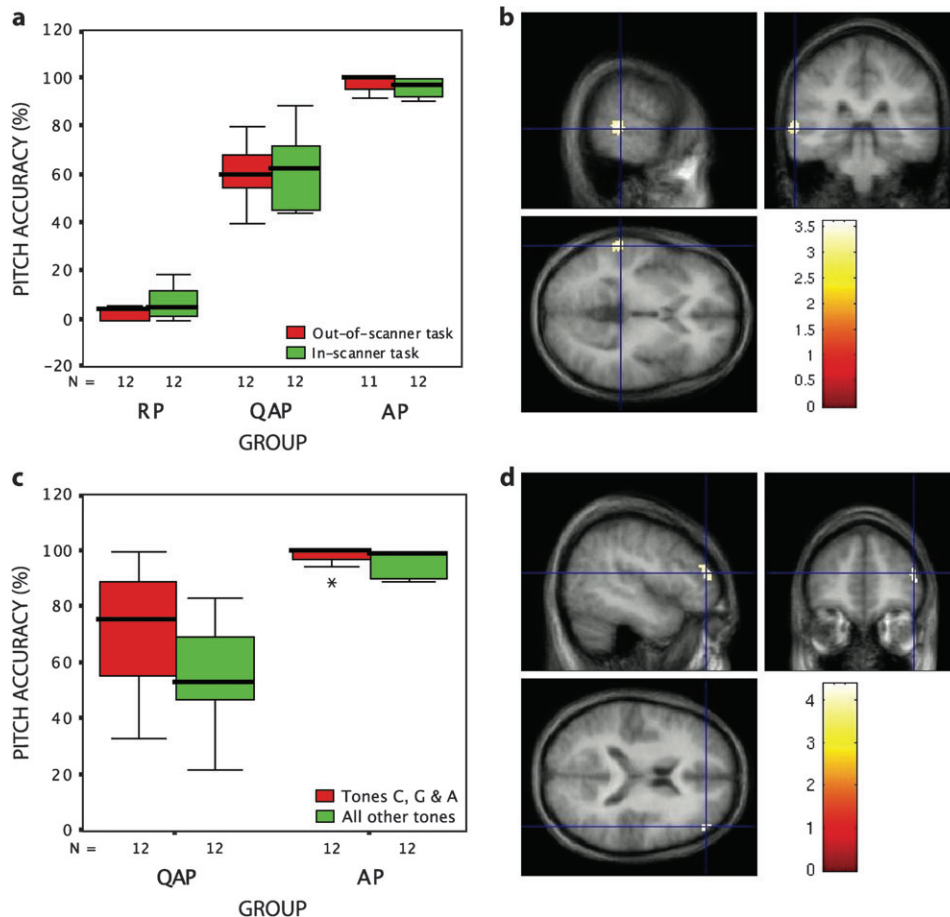
Figure 2. Pitch naming. (*a*) Box and whisker plots of the accuracy of pitch naming, shown as percentage (%) correct for the out-of-scanner (red) and in-scanner (green) behavioral tasks. Note that in-scanner performance fell within the appropriate range for each group classified from the out-of-scanner task. (*b*) A statistical parametric map of the contrast (pitch naming—baseline) for the ROI analysis over the left superior temporal gyrus in the AP group, overlaid on the averaged normalized brain of these musicians. The threshold was set at $P < 0.05$ corrected, with peak voxels indicating activation in the more posterior extent of the left superior temporal gyrus (BA 22). (*c*) Response accuracy, shown as percentage (%) correct for the notes C, G, and A (red) compared to all other notes (green) for the AP and QAP groups. (*d*) A statistical parametric map of the contrast (pitch naming—tonal classification) for the ROI analysis over the right middle frontal gyrus in the QAP group, overlaid on the averaged normalized brain of these musicians. The threshold was set at $P < 0.05$ corrected, with peak voxels indicating right hemisphere activation in the region of BA 46.

In order to investigate our first aim, we used the subtraction method to identify changes in cerebral blood flow associated with explicit pitch naming minus baseline in the traditionally high performing AP group. In accordance with our first prediction, there was a significant activation peak in the posterior extent of the left superior temporal gyrus (Brodmann area [BA] 22) in the AP group during explicit pitch naming (–61, –31, 5; $Z = 3.56$, $P < 0.05$ with small volume correction for superior temporal gyrus ROI; Fig. 2*b*). The anatomical location of this activation peak lay within the superior temporal gyrus of 10 of the AP musicians and within the middle temporal gyrus in the remaining 2. This finding provides support for the role of this region in an AP template.

The Influence of Expertise on Pitch Naming

The QAP group showed lower accuracy and greater variability in their pitch naming performance (Fig. 2*a*) however there was no difference in their average response times for correctly identified notes compared with AP musicians during scanning (AP = 2,150 ± 167 [SD] ms; QAP = 2,352 ± 386 ms, $t_{22} = 1.660$, $P > 0.05$). This suggests a similar level of response automaticity for tones that were absolutely encoded by the AP and QAP

musicians, albeit for fewer tones or a more limited AP template in the QAP musicians.

To address our second question, we initially examined cerebral blood flow changes for explicit pitch naming minus baseline in the combined AP and QAP groups. This reflected the lack of difference in reaction times of AP and QAP musicians for correctly identified tones, and thus the possibility of similar AP representation in the 2 groups. The analysis showed that in addition to increased blood flow in the left superior temporal gyrus, as observed in the AP group alone, there was activation of an extensive right hemisphere network (see Table 2), including bilateral activation of frontotemporal regions previously implicated in pitch discrimination and auditory working memory functions (Zatorre et al. 1994). In other words in the combined AP and QAP group, including musicians with lower AP skill (QAP musicians) appeared to recruit more extensive neural networks to support pitch naming performance, producing a qualitatively different pattern to that observed for the AP group alone. The influence of expertise on the extent of activation is consistent with previous research investigating musical expertise using functional neuroimaging (Münte et al. 2002).

Variability in the AP Template

Variability in the AP template was highlighted by the self-reported strategies of QAP musicians to facilitate pitch naming performance. These strategies showed context dependence for pitch cues and naturally fell into 2 categories that were not mutually exclusive, namely: 1) timbral facilitation, characterized by superior pitch naming for certain instrumental timbres often in a restricted pitch range, and 2) the use of reference tones, with the musical notes C, G, and A identified as salient (see Table 3). This latter strategy engages RP processing, with the reference tone serving as an anchor from which interval judgments are made to identify remaining tones. Approximately half (42%) of the QAP musicians reported regular use of reference tones, and in support of this claim, QAP musicians showed more accurate identification of the notes C, G, and A compared with all other notes during scanning ($t_{11} = 2.474$, $P < 0.05$), which was not evident for the AP group (Fig. 2c).

To further investigate our second question and in particular, to minimize the influence of RP judgments on cerebral blood flow changes in QAP musicians, we subtracted tonal classification from explicit pitch naming in the QAP group alone. The

results again revealed activation of an extensive right hemisphere network in the QAP musicians, including peaks along the anterior–posterior extent of the right superior temporal gyrus, right middle and inferior frontal gyri, and right cerebellum (see Table 4). Maximal blood flow change was observed in the right middle frontal gyrus in the region of BA 46 ($48, 45, 16$; $Z = 4.29$, voxel-level corrected $P < 0.01$, see Fig. 2d). This suggests that a more limited AP template requires musicians to hold tones "in mind" when performing AP judgments. In other words, the engagement of pitch working memory structures at relatively high processing load provides a potential neurocognitive marker of lower AP skill (cf. Münte et al. 2002; Ross et al. 2004).

Differentiating the Neurocognitive Components of Pitch Naming

There were some striking behavioral differences between the in-scanner pitch naming performance of QAP musicians who primarily endorsed a reference tone strategy compared with timbral facilitation, despite the small sample sizes of these subgroups (see Table 3). The reference tone subgroup mislabeled more tones than the timbral subgroup ($t_8 = -2.385$, $P < 0.05$), particularly tones corresponding to the black keys of the piano ($t_{5.048} = -3.254$, $P < 0.05$) (Fig. 3a). Both subgroups showed similar response accuracy for the notes C, G, and A, however the timbral subgroup showed greater accuracy for all remaining tones ($t_8 = 2.611$, $P < 0.05$) (Fig. 3b), with the reference tone subgroup making significantly more semitone errors (reference tone mean 16.4 ± 7.2 errors; Timbral mean 7.4 ± 4.6 errors, $t_8 = -2.346$, $P < 0.05$). Taken together, these findings suggest a qualitative difference between the 2 subgroups that reflects skill expertise, with greater accuracy of pitch naming associated with a more extensive AP template, and presumably less need to employ pitch working memory strategies.

To investigate this notion further, an "AP automatic template" group was formed comprising musicians in the traditionally high AP group and those reporting a timbral or automatic pitch processing strategy in the QAP group (see Table 3). Examination of cerebral blood flow changes associated with pitch naming minus baseline in the AP automatic template group confirmed the same significant activation peak in the posterior extent of the left superior temporal gyrus (-61, -31, 5) with no associated changes in right frontal pitch working memory structures.

RP Processing

For the in-scanner tonal classification task mean performance accuracy (scored out of 54) was similarly high for the AP, QAP, and RP groups, supporting the utility of this RP functional

---

Table 2

Right hemisphere network and left frontotemporal regions activated during pitch naming in the AP and QAP musicians

Region	BA	x	y	z
R inferior temporal gyrus*	20	55	−55	−12
R middle frontal gyrus	46	51	34	20
R anterior cingulate	32	14	28	18
R cerebellum		4	−73	−20
		26	−59	−24
L superior temporal gyrus (posterior)	22	−59	−33	3
L superior temporal gyrus (anterior)	22	−57	−10	−1
L middle frontal gyrus	11	−34	43	−4
L inferior frontal gyrus	11	−24	36	−20
	45	−50	17	19

Note: Regions of increased cerebral blood flow are shown for voxel-level uncorrected $P < 0.001$.
*Voxel-level corrected $P < 0.05$.

Table 3

Self-reported pitch naming strategies of the QAP musicians

Subgroup	Self-reported strategy
Timbral	Best for familiar instrumental timbres in the middle range
Timbral	Best for the white notes of the piano
Timbral	Best for B flat trumpet range, uses RP when "lost"
Timbral	Best for the white notes of the piano within the musician's vocal range
Timbral	Best for the white notes of the piano
Automatic	"I have a natural keyboard programmed in my mind"
Automatic	Improvement with repetition or increased response time
Reference	Uses "C," "G," and "A" or the pitch of a known song
Reference	Uses "C"; best for the middle range of certain instruments, finds brass timbres difficult
Reference	Uses "A" and "C"; best for piano timbre; more variable for other instruments
Reference	Uses "C"; performance is facilitated by key membership
Reference	Compares pitch to a known song if pitch name is not automatically present; best for piano tones in the upper range, visualizes fingers on the keyboard

Note: All QAP musicians described automatic responses for the identification of some notes, and similar to the AP musicians, 2 QAP participants reported an automatic response for all notes. More commonly however, the QAP musicians reported cognitive strategies that facilitated their pitch naming performance, with those musicians reporting regular use of a reference tone placed in the reference tone subgroup regardless of other strategies reported. The 2 musicians reporting automatic coding only were excluded from the reference tone versus timbral subgroup comparisons described in the text.

Table 4

Right hemisphere network activated during pitch naming after subtraction of RP processing in the QAP musicians

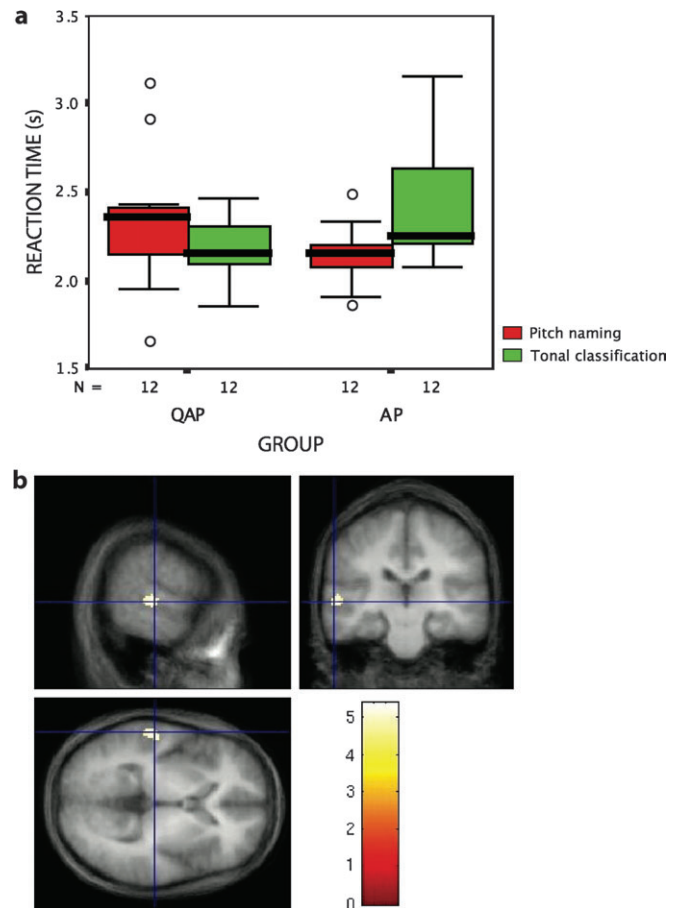| Right hemisphere region	BA	x	y	z
Middle frontal gyrus	46	48	45	16
Inferior frontal gyrus	45	59	20	17
Superior temporal gyrus (anterior)	22	59	7	−5
Superior temporal gyrus (posterior)	22	67	−37	4
	39	59	−58	8
Right cerebellum		51	−67	−25

Note: Regions of increased cerebral blood flow are shown for voxel-level uncorrected $P < 0.001$.
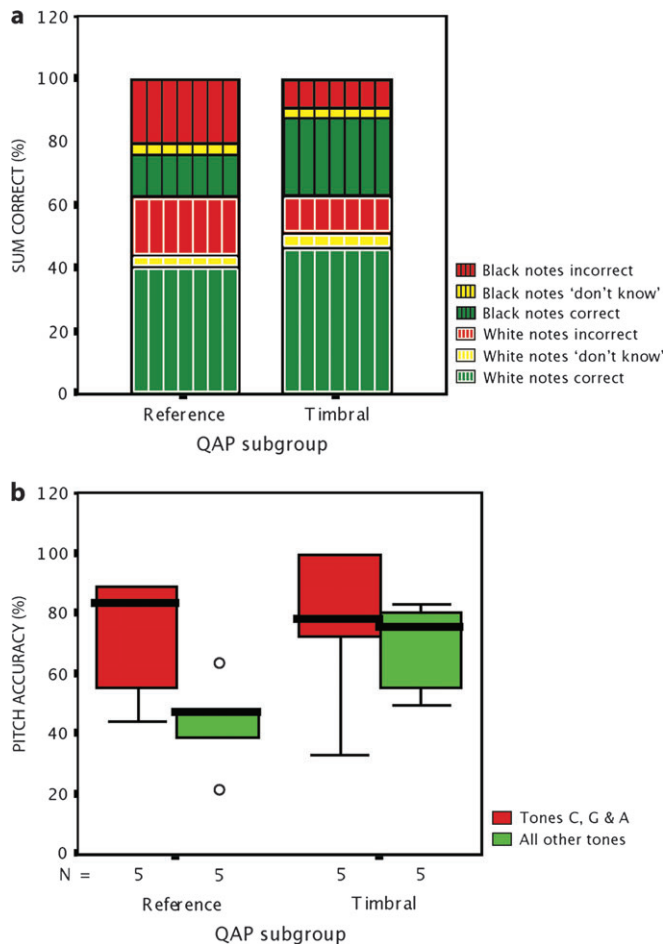
Figure 3. Variability in the AP template. (*a*) Response accuracy of pitch naming, shown as percentage (%) correct for tones corresponding to the white and black notes of the piano for the timbral and reference tone subgroups of the QAP musicians. (*b*) Box and whisker plots of the accuracy of pitch naming, shown as percentage (%) correct for the notes C, G, and A (red) compared to all other notes (green) for the timbral and reference tone subgroups of the QAP musicians.

Figure 4. RP processing. (*a*) Box and whisker plots of the response times of the AP and QAP groups for correctly identified items for the pitch naming (red) and tonal classification (green) tasks. (*b*) A statistical parametric map of the contrast [tonal classification—baseline] for musicians in the RP and reference tone groups, overlaid on the averaged normalized brain of these musicians. The threshold was set at $P <$ 0.05 corrected, with peak voxels indicating more anterior activation in the left superior temporal gyrus relative to pitch naming (Fig. 2*b*).

imaging task (AP = 50.67 ± 6.33; QAP = 50.83 ± 6.75; RP = 53.42 ± 1.24, $F_{2,33}$ = 0.981, $P >$ 0.05). Of note, QAP musicians showed significantly faster mean response time for correct tonal classification compared with correct pitch naming, whereas the reverse was true for AP musicians ($F_{1,22}$ = 9.093, $P <$ 0.01, see Fig. 4*a*). Some AP musicians spontaneously reported mentally translating tones from their pitch names to their RP classification during scanning, likely accounting for the slower mean reaction time of AP musicians for correct tonal classification compared with QAP musicians (contrast estimate = –0.243, $P =$ 0.02; Fig. 4*a*). There was no difference between the mean reaction times of the QAP and RP musicians (RP = 2,290 ± 171 ms, contrast estimate = 0.109, $P >$ 0.05).

Visual inspection of the data revealed similar patterns of activation for the AP, QAP, and RP groups for the tonal classification task. Thus, we performed 2 contrasts to investigate the neurocognitive basis of the performance disadvantage of the AP musicians on the RP processing task. The first identified increased blood flow associated with RP processing common to all musicians, by subtracting the baseline condition from the tonal classification task in the AP, QAP, and RP groups. This revealed significant peaks of activation in the left superior temporal gyrus (–57, –25, 5; Z = 5.60; –57, –17, 5; Z = 5.19; –50, 6, –2; Z = 4.97, all voxel-level corrected $P <$ 0.01), and the right cerebellum (4, –75, –16; Z = 4.47, voxel-level corrected $P <$ 0.05). The peaks were more anterior along the extent of the left superior temporal gyrus in the sagittal plane, but proximal to the significant peak observed during pitch naming minus baseline (Fig. 2*b*). The second contrast assessed the presence of any differences associated with RP processing in the absence of automatic pitch naming. Thus, musicians assigned to the AP automatic template group were removed from the subtraction of tonal classification minus baseline. This showed a sole region of increased cerebral blood flow at the same more anterior peak within the left superior temporal gyrus (–57, –25, 5; Z = 5.23, voxel-level corrected $P <$ 0.05, see Fig. 4*b*) for the musicians principally using RP processing. Taken together, one interpretation of these findings is that an RP performance disadvantage in AP musicians reflects automatic engagement of the AP template during RP processing, with an associated decrease in RP processing speed.

Structural Differences in the Planum Temporale
To investigate the presence of structural differences in the planum temporale of the AP, QAP, and RP musicians, a left-right
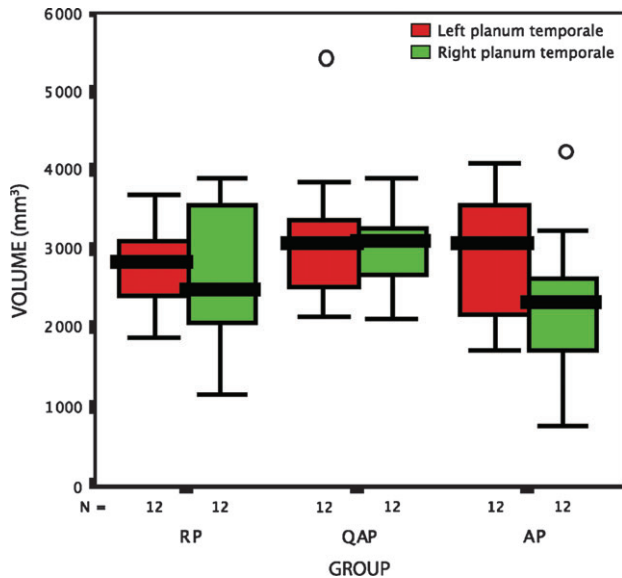
Figure 5. Box and whisker plots of the left (red) and right (green) planum temporale volumes of the AP, QAP, and RP musicians. Note that the 2 outliers were excluded from the statistical analyses reported in the text.

asymmetry index was calculated for each group to correct for differences in total planum temporale size (Steinmetz 1996). The AP musicians showed the greatest leftward asymmetry (AP = −0.306 ± 0.321; QAP = −0.002 ± 0.264; RP = −0.088 ± 0.364) that was due to a significantly smaller mean right planum temporale volume compared with the QAP and RP musicians ($F_{2,31}$ = 3.834, P < 0.05, see Fig. 5). Planned simple contrasts revealed that this difference was most evident between the AP and QAP musicians (contrast estimate = 815.09, P = 0.011), with the QAP musicians failing to show the typical leftward asymmetry of the planum temporale (t_{10} = 0.082, P > 0.05; Fig. 5). There was no difference between the mean left planum temporale volumes of the AP, QAP, and RP groups ($F_{2,31}$ = 0.153, P > 0.05).

The smaller mean right planum temporale volume of the AP group did not appear to reflect general plasticity effects associated with an earlier age of onset of musical training. Planned difference contrasts revealed earlier training onset in both the AP and QAP musicians compared with the RP musicians (AP = 4.3 ± 0.9 years; QAP = 5.6 ± 1.9 years; RP = 6.8 ± 1.9 years, $F_{2,33}$ = 6.77, contrast estimate = 1.812, P = 0.004), whereas there was no difference between the AP and QAP groups (contrast estimate = 1.292, P > 0.05). A significant negative correlation was also evident between age of onset of musical training and performance of the in-scanner AP task (Pearson r = −0.572, P < 0.001), with the distribution of scores indicating high response accuracy (>80%) only in musicians who commenced training between 2 and 6 years. These findings support the notion of a sensitive period for the development of AP.

Discussion

In this study we combined structural volumetric and functional cerebral blood flow measures with behavioral measures within and outside the scanner in musicians with varying AP skill that point to involvement of the left planum temporale in pitch naming ability. Significant activation associated with this region

appeared dependent on high levels of skill expertise, with less skilled performance engaging a right hemisphere network including pitch working memory structures. The automaticity of pitch naming conferred a disadvantage on RP processing speed that may be underpinned by activation of proximal regions in the left superior temporal gyrus. The functional imaging findings were supported by structural asymmetry of the planum temporale that was most evident in the AP musicians despite an early age of onset of musical training in both the AP and QAP groups.

We believe the findings highlight the utility of investigating individuals with partial representations to gain insight into the neurobiological complexities underpinning knowledge formation. Our goal was to investigate differences in AP ability, and thus we chose methods that allowed us to probe these differences (i.e., use of a piano timbre to examine pitch naming in AP and QAP musicians). Previous research has generally excluded individuals with QAP or used sine tones to examine pitch naming, in effect removing variability that naturally exists in the population (see Athos et al. 2007). Arguably, this has led to a restricted understanding of the mechanisms underpinning AP skill, to which we believe the current study sheds important insights. In particular, by systematically examining variable pitch naming ability we have identified component neurocognitive processes that are differentially involved in the representation of pitch. These processes are underscored by changes in brain structure and function that were reflected by our in-scanner behavioral performance measures. Conceivably these changes are the expression of complex interactions between early learning experiences and genetic predispositions (Zatorre 2003).

First, high pitch naming accuracy associated with faster response times for AP over RP judgments showed principal involvement of the left hemisphere, indicated by a more posterior peak along the anterior–posterior extent of the left superior temporal gyrus. Our in-scanner pitch naming task was purposely designed to remove cerebral blood flow changes previously attributed to conditional associative and auditory working memory processes in frontal regions of the brain. This allowed examination of the functional role of the left posterior superior temporal gyrus in the long-term representation of pitch. Interestingly, in the majority of AP musicians peak activation fell on the bank of the superior temporal gyrus, abutting the superior temporal sulcus. The superior temporal sulcus has been generally implicated in the integration of multimodal sensory input (Noppeney et al. 2007) and more specifically, in different levels of pitch pattern analysis that may be relevant to linguistic prosody (Stewart et al. 2008). Thus our findings are consistent with the notion that high pitch naming accuracy may be supported by a unique pitch template, where the verbal code is integrated with the AP representation (Levitin and Rogers 2005). In our study, the decreased volume of the right planum temporale of AP musicians is consistent with activation of the left posterior superior temporal gyrus during pitch naming (see also Keenan et al. 2001). It implies a reduced role for the homologous region in the right hemisphere in individuals with a high degree of skill automaticity. This automaticity conferred a RP processing disadvantage in AP musicians, with engagement of pitch naming requiring translation for successful RP judgments in some AP musicians. This may account for the slower RP response times of the AP musicians.

Second, a more limited AP template was associated with more extensive recruitment of neuronal networks, particularly in the right hemisphere. This appeared commensurate with the extent to which QAP musicians used auditory working memory strategies. The use of reference tones by QAP musicians and the salience of tones C, G, and A, have been previously well documented (Miyazaki 1990; Takeuchi and Hulse 1993; Athos et al. 2007). Research has also shown engagement of pitch working memory processes by less expert musicians associated with activation of right dorsolateral prefrontal cortex (Zatorre et al. 1994). The use of pitch working memory has recently been proposed as a marker of lower AP skill (Ross et al. 2004). The current study brings these findings together by demonstrating peak activation in right dorsolateral prefrontal cortex in QAP musicians principally employing a reference tone strategy.

In addition to the use of reference tones, the behavioral data support previous findings that QAP is facilitated by a range of auditory cues, including timbre, key color, and pitch register or height (Bachem 1955; Takeuchi and Hulse 1993, Athos et al. 2007). In the present study, this may have been associated with activation of the right superior temporal gyrus previously shown to be important in fine-grained spectral processing (Griffiths and Warren 2002, Stewart et al. 2006; Griffiths et al. 2007). In particular, activation of the more posterior extent of the right superior temporal gyrus in QAP musicians suggests involvement of this homologous region in pitch representation. This is consistent with the significantly larger mean right planum temporale volume of the QAP musicians. This finding is intriguing, and warrants further investigation. One possibility is that QAP reflects a more limited AP template that is more reliant on the presence of contextual cues during perceptual encoding. The nature of these cues may reflect experiences from early musical training, including exposure to white notes before the black notes of the piano, typically in a restricted pitch range. In other words, general plasticity effects associated with an early age of musical training per se cannot account for the volume differences observed in AP and QAP musicians. Rather, a more limited AP template appears dependent on contextual cues that may be present during early training and may be associated with structural differences in the organization of long-term pitch representation.

Third, we found activation of the more anterior extent of the left superior temporal gyrus during RP processing in musicians with or without AP. Previous research has implicated this region in RP processing including the detection of violations of tonality (Janata et al. 2002, Warren et al. 2003), with tracking of tonal space maintained by regions in the prefrontal cortex (Janata et al. 2002). Our tonal classification task required identification of intervals conforming to the Western tonal idiom after subtraction of auditory working memory processes, and thus our findings are in keeping with this previous research. They also contribute new data on the possibility of a proximal, dual representation of absolute and RP in AP possessors, potentially accounting for the RP performance disadvantage of AP musicians and previous suggestions of an integrated representation of AP and interval information (Levitin and Rogers 2005).

In summary, the study findings are based on converging effects from a number of methodological approaches. There were clear differences in the volumes of the right planum temporale of the musicians with AP, QAP, and RP. These differences were accompanied by different patterns of functional activation observed for the AP and QAP musicians on the pitch naming task. The AP musicians showed a smaller right planum temporale and unilateral left-sided activation in the more posterior extent of the superior temporal gyrus during pitch naming. In contrast, the QAP musicians showed symmetry of the planum temporale and greater activation of the right superior temporal gyrus during pitch naming. These effects were associated with clear differences in task performance at a behavioral level. Taken together, one explanation of these congruent findings is plasticity of auditory knowledge representation in the context of environmental fostering of early cognitive development that translates to differences in cognitive ability. Arguably this involves long-term cortical changes at a system level impacting brain structure, function, and behavior.

Conclusions

This study provides new insights into the spectrum of neurocognitive processes that may underpin absolute and RP processing. Central to our findings is the establishment of a neurobiological basis for an AP template in the temporal lobe. More broadly, our findings illustrate the importance of systematically examining variability in musical skill expertise to provide an integrated account of pitch processing in the brain. This account informs our understanding of the development of knowledge and cognitive skills and their cerebral representation.

References

Athos EA, Levinson B, Kistler A, Zemansky J, Bostrom A, Freimer N, Gitschier J. 2007. Dichotomy and perceptual distortions in absolute pitch ability. Proc Natl Acad Sci USA. 104:14795–14800.

Bachem A. 1955. Absolute pitch. J Acoust Soc Am. 27:1180–1185.

Baharloo S, Johnston PA, Service SK, Gitschier J, Freimer NB. 1998. Absolute pitch: an approach for identification of genetic and nongenetic components. Am J Hum Genet. 62:224–231.

Baharloo S, Service SK, Risch N, Gitschier J, Freimer NB. 2000. Familial aggregation of absolute pitch. Am J Hum Genet. 67:755–758.

Bermudez P, Zatorre RJ. 2005. Conditional associative memory for musical stimuli in nonmusicians: implications for absolute pitch. J Neurosci. 25:7718–7723.

Deutsch D. 2002. The puzzle of absolute pitch. Curr Dir Psychol Sci. 11:200–204.

Gregersen PK, Kowalsky E, Kohn N, Marvin EW. 2001. Early childhood music education and predisposition to absolute pitch: teasing apart genes and environment. Am J Med Genet. 98:280–282.

Griffiths TD, Kumar S, Warren JD, Stewart L, Enno Stephan K, Friston KJ. 2007. Approaches to the cortical analysis of auditory objects. Hear Res. 229:46–53.

Griffiths TD, Warren JD. 2002. The planum temporale as a computational hub. Trends Neurosci. 25:348-353.

Halpern AR. 1989. Memory for the absolute pitch of familiar songs. Mem Cogn. 17:572-581.

Janata P, Birk J, Van Horn JD, Leman M, Tillmann B, Bharucha JJ. 2002. The cortical topography of tonal structures underlying Western music. Science. 298:2167-2170.

Keenan JP, Thangaraj V, Halpern AR, Schlaug G. 2001. Absolute pitch and planum temporale. Neuroimage. 14:1402-1408.

Krumhansl CL. 1991. Music psychology: tonal structures in perception and memory. Annu Rev Psychol. 42:277-303.

Lancaster JL, Woldorff MG, Parsons LM, Liotti M, Freitas CS, Rainey L, Kochunov PV, Nickerson D, Mikiten SA, Fox PT. 2000. Automated Talairach atlas labels for functional brain mapping. Hum Brain Mapp. 10:120-131.

Levitin DJ. 1994. Absolute memory for musical pitch: evidence from the production of learned melodies. Percept Psychophys. 56: 414-423.

Levitin DJ, Rogers SE. 2005. Absolute pitch: perception, coding, and controversies. Trends Cogn Sci. 9:26-33.

Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH. 2003. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. Neuroimage. 19:1233-1239 (WFU Pickatlas, version 2.0).

Miyazaki K. 1990. The speed of musical pitch identification by absolute-pitch possessors. Mus Percept. 8:177-188.

Miyazaki K, Rakowski A. 2002. Recognition of notated melodies by possessors and non-possessors of perfect pitch. Percept Psychophys. 64:1337-1345.

Münte TF, Altenmüller E, Jäncke L. 2002. The musician's brain as a model of neuroplasticity. Nat Rev Neurosci. 3:473-478.

Noppeney U, Josephs O, Hocking J, Price CJ, Friston KJ. 2007. The effect of prior visual information on recognition of speech and sounds. Cereb Cortex. doi:10.1093/cercor/bhm091.

Ohnishi T, Matsuda H, Asada T, Aruga M, Hirakata M, Nishikawa M, Katoh A, Imabayashi E. 2001. Functional anatomy of musical perception in musicians. Cereb Cortex. 11:754-760.

Profita J, Bidder GT. 1988. Perfect pitch. Am J Med Genet. 29: 763-771.

Ross DA, Olson IR, Marks LE, Gore JC. 2004. A non-musical paradigm for identifying absolute pitch possessors. J Acoust Soc Am. 116: 1793-1799.

Saffran JR, Griepentrog GJ. 2001. Absolute pitch in infant auditory learning: evidence for developmental reorganization. Dev Psychol. 37:74-85.

Schlaug G, Jäncke L, Huang Y, Steinmetz H. 1995. In vivo evidence of structural brain asymmetry in musicians. Science. 267:699-701.

Steinmetz H. 1996. Structure, function and cerebral asymmetry: in vivo morphometry of the planum temporale. Neurosci Biobehav Rev. 20:587-591.

Steinmetz H, Rademacher J, Huang YX, Hefter H, Zilles K, Thron A, Freund HJ. 1989. Cerebral asymmetry—MR planimetry of the human planum temporale. J Comput Assist Tomogr. 13:996-1005.

Stewart L, Overath T, Warren JD, Foxton JM, Griffiths TD. 2008. fMRI evidence for a cortical hierarchy of pitch pattern processing. PloS ONE. 3:e1470. doi:10.1371/journal.pone.0001470.

Stewart L, von Kriegstein K, Warren JD, Griffiths TD. 2006. Music and the brain: disorders of musical listening. Brain. 129:2533-2553.

Takeuchi AH, Hulse SH. 1993. Absolute pitch. Psychol Bull. 113: 345-361.

Talairach J, Tourneaux P. 1988. Co-planar stereotaxic atlas of the human brain. New York: Thieme Medical.

Terhardt E, Ward WD. 1982. Recognition of musical key: exploratory study. J Acoust Soc Am. 72:26-33.

Trehub SE. 2003. Toward a developmental psychology of music. Ann N Y Acad Sci. 999:402-413.

Trehub SE, Hannon EE. 2006. Infant music perception: domain-general or domain-specific mechanisms? Cognition. 100:73-99.

Warren JD, Uppenkamp S, Patterson RD, Griffiths TD. 2003. Separating pitch chroma and pitch height in the human brain. Proc Natl Acad Sci USA. 100:10038-10042.

Woods RP, Mazziotta JC, Cherry SR. 1993. MRI-PET registration with automated algorithm. J Comput Assist Tomogr. 17:536-546.

Worsley KJ, Evans AC, 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. 2003. Absolute pitch: a model for understanding the influence of genes and development on neural and cognitive function. Nat Neurosci. 6:692-695.

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

Zatorre RJ, Perry DW, Beckett CA, Westbury CF, Evans AC. 1998. Functional anatomy of musical processing in listeners with absolute pitch and relative pitch. Proc Natl Acad Sci USA. 95:3172-3177.