◆ **Human Brain Mapping 000:00–00 (2012)** ◆

Auditory Stroop and Absolute Pitch: An fMRI Study

Katrin Schulze,[1,2]* **Karsten Mueller,**[1] **and Stefan Koelsch**[1,3]

[1]*Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany*
[2]*Developmental Cognitive Neuroscience Unit, UCL Institute of Child Health, London, United Kingdom*
[3]*Cluster of Excellence "Languages of Emotion," Freie Universität Berlin, Berlin, Germany*

◆

Abstract: To date, the underlying cognitive and neural mechanisms of absolute pitch (AP) have remained elusive. In the present fMRI study, we investigated verbal and tonal perception and working memory in musicians with and without absolute pitch. Stimuli were sine wave tones and syllables (names of the scale tones) presented simultaneously. Participants listened to sequences of five stimuli, and then rehearsed internally either the syllables or the tones. Finally participants indicated whether a test stimulus had been presented during the sequence. For an auditory stroop task, half of the tonal sequences were congruent (frequencies of tones corresponded to syllables which were the names of the scale tones) and half were incongruent (frequencies of tones did not correspond to syllables). Results indicate that first, verbal and tonal perception overlap strongly in the left superior temporal gyrus/sulcus (STG/STS) in AP musicians only. Second, AP is associated with the categorical perception of tones. Third, the left STG/STS is activated in AP musicians only for the detection of verbal-tonal incongruencies in the auditory stroop task. Finally, verbal labelling of tones in AP musicians seems to be automatic. Overall, a unique feature of AP appears to be the similarity between verbal and tonal perception. Hum Brain Mapp 00:000–000, 2012. © 2012 Wiley Periodicals, Inc.

Key words: absolute pitch; auditory perception; auditory working memory; auditory stroop; superior temporal gyrus/sulcus

◆

INTRODUCTION

Absolute pitch (AP) is defined as the ability to identify any pitch of the Western musical scale without an external reference tone [Miyazaki, 1988; Takeuchi and Hulse, 1993]. To date, the underlying cognitive and neural mechanisms of AP have remained elusive. Although AP is a rare ability, understanding the underlying processes is fundamental because AP can serve as a model to understand how very special abilities are represented in the brain [Zatorre, 2003] and to investigate general human perceptual and mnemonic processes in the auditory domain [Bermudez and Zatorre, 2009]. Different hypotheses have been put forward to explain AP, for example, categorical perception/processing of pitch information [Rakowski, 1993; Siegel, 1974], the association of tones with verbal labels [Siegel, 1974], the use of multiple coding strategies [Zatorre and Beckett, 1989], and the assumption of internal templates of tones which decrease the working memory (WM) load for tonal material [Hantz et al., 1992; Klein et al., 1984; Zatorre et al., 1998]. These different hypotheses are not necessarily mutually exclusive. Zatorre [2003] acknowledged the potential complexity of the underlying mechanisms of AP by ascribing AP to two separable

---

Additional Supporting Information may be found in the online version of this article.

*Correspondence to: Katrin Schulze, Developmental Cognitive Neuroscience Unit, UCL Institute of Child Health, 30 Guilford Street, London WC1N 1EH, United Kingdom.
E-mail: kschulze@ich.ucl.ac.uk

Received for publication 9 May 2011; Revised 26 September 2011; Accepted 15 November 2011

DOI: 10.1002/hbm.22010

cognitive components. He suggested first, that AP possessors perceive and encode tones within very narrow and fixed pitch categories (first component: perception/encoding), and second, that these categorized pitches are subsequently associated with (verbal) labels (associative memory).

First Component—Perception/Encoding

Several studies [Rakowski, 1993; Siegel, 1974] suggested that AP musicians perceive tones categorically, in contrast to individuals without AP. In line with this hypothesis, neurophysiological differences were observed between AP and non-AP musicians during the perception of tones [Hirata et al., 1999; Itoh et al., 2005; Wu et al., 2008], suggesting that perception and early auditory encoding of tones might rely on different processes in AP compared to non-AP musicians. Using fMRI, Schulze et al. [2009] observed a stronger activation of the left superior temporal sulcus (STS) in AP musicians compared to non-AP musicians during tone perception. Independent of the ability of AP, an increased activation of the STS is typically found when the perception of speech is compared with the perception of non-speech in humans [Binder et al., 2000; Dehaene-Lambertz et al., 2005; Jancke et al., 2002; Liebenthal et al., 2005; Specht et al., 2005]. This difference is presumably due to a categorical (e.g., speech) versus a non-categorical (e.g., non-speech) perception, which involves the STS and/or the superior temporal gyrus [STG; Leech et al., 2009; Liebenthal et al., 2005; Luo et al., 2005; Mottonen et al., 2006]. Therefore, the stronger activation of the left STS in AP compared to non-AP musicians during pitch perception indicates a more categorical perception in AP musicians [Bermudez and Zatorre, 2009; Schulze et al., 2009]. That is, the neural correlates subserving verbal and tonal perception might show a larger overlap in musicians with than without AP.

At a structural brain level, it has been suggested that AP is associated with an increased leftward planum temporale asymmetry [Keenan et al., 2001; Luders et al., 2004; Schlaug et al., 1995; Zatorre et al., 1998]; however, these results remain inconsistent [see Bermudez et al., 2009]. Other studies also indicated that AP musicians show a changed connectivity in the temporal lobe [Loui et al., 2010; Oechslin, Imfeld et al., 2009]. Therefore, accumulating data indicate that AP is associated with both structural [Keenan et al., 2001; Loui et al., 2010; Luders et al., 2004; Oechslin, Imfeld et al., 2009; Schlaug et al., 1995; Zatorre et al., 1998] and functional [Ohnishi et al., 2001; Schulze et al., 2009] modifications in the temporal lobe.

Second Component—Associative Memory

The association of tones with verbal labels [Siegel, 1974; Takeuchi and Hulse, 1993] or multiple codes [Zatorre and Beckett, 1989] has been suggested to play a role in AP and is thought to be reflected functionally and anatomically in the posterior dorsal frontal region [Bermudez et al., 2009; Bermudez and Zatorre, 2005; Zatorre et al., 1998]. In the search for the underlying cognitive neural substrate of AP, short-term memory and working memory had also been investigated. Event-related potential (ERP) studies indicate that AP musicians may not, or to a lesser degree, require a WM update during a pitch memory task compared with non-AP musicians [Hantz et al., 1992; Itoh et al., 2005; Klein et al., 1984; Wayman et al., 1992; but see also Hirose et al., 2002], because internal templates facilitate memory for tones [Zatorre, 2003]. These findings were corroborated by fMRI results showing more activation in the right parietal lobe (superior parietal lobule/intraparietal sulcus; SPL/IPS) in non-AP compared to AP musicians during a pitch WM task [Schulze et al., 2009], an area known to be involved in WM [Baddeley, 2003; Schulze et al., 2011a].

To acknowledge potential differences between encoding/perception and WM [Bermudez and Zatorre, 2009; Schulze et al., 2009; Zatorre, 2003] we investigated these processes separately in the present fMRI study. Tonal and verbal perception and WM rehearsal were compared between AP musicians and non-AP musicians, to our knowledge for the first time. If, as suggested, verbal and tonal perception are more similar in musicians with than without AP [Schulze et al., 2009], then a difference should be observed in the left STS in non-AP musicians when comparing verbal versus tonal perception, but not in the AP musicians. During tonal WM stronger activation of classical WM areas (Broca's area, left premotor cortex, left parietal lobe, and cerebellum) in non-AP musicians, compared to AP musicians [Schulze et al., 2009] was expected.

Auditory Stroop Task

Miyazaki [2004] conducted an 'auditory stroop' experiment by presenting AP musicians with congruent (sung tone name and frequency of sung tone name were the same) and incongruent auditory stimuli (sung tone name and frequency of sung tone name were different). Participants were asked to repeat the syllables, ignoring the pitch. The results showed longer reaction times for incongruent than congruent stimuli, supporting the notion that verbal labeling of tones is automatic for AP musicians and difficult to suppress, even if disadvantageous. These results were replicated by Itoh et al. [2005], who in addition demonstrated that a left posterior temporal negativity in AP possessors ("AP negativity") with a latency of 150 ms was modulated by the incongruency during a listening and pitch-naming task. In the present experiment we investigated the neural correlates of perception and WM of congruent and incongruent sequences in musicians with and without AP. In contrast to non-AP musicians, AP musicians were expected to show a superior WM performance for the congruent compared to the incongruent stimuli [Itoh et al., 2005; Miyazaki, 2004; Siegel, 1974].

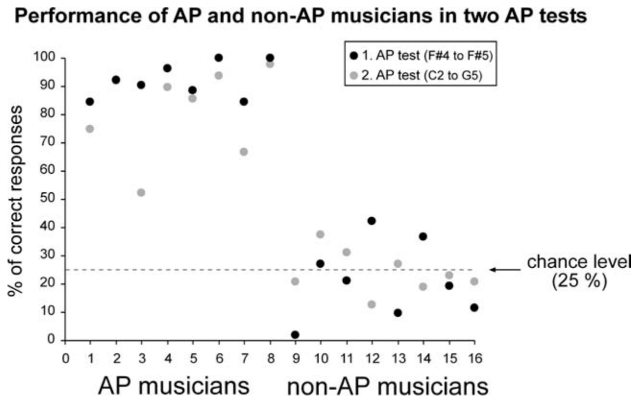

Figure 1.

Performance in two AP test for AP and non-AP musicians. Chance level was 25% (answers within one semitone of the presented pitch were regarded as a correct answer; see Materials and Methods).

Furthermore, areas in the left temporal lobe [Itoh et al., 2005] were hypothesized to show a difference in AP, but not in non-AP, musicians between the perception of congruent and incongruent sequences.

MATERIALS AND METHODS

Stimuli, procedure, and scanning paradigm have been employed by our group in previous experiments [Schulze et al., 2011a,b].

Participants

AP and non-AP musicians were recruited from the "University of Music and Theatre Mendelsohn Bartholdy" in Leipzig. Eight musicians (three males) were assigned to the AP group after performing two AP tests (see below). Within the AP group mean age was 25.4 years (SD = 4.57; range: 21–33). AP musicians had started their musical training at an average age of 5.94 years (SD = 0.86 years) and their AP ability was discovered on average when they were 9 years old (SD = 5.32). Seven of the AP musicians were not native German speakers, but they studied in Germany and spoke German fluently (and, as can be seen in the results, no difference in the verbal WM task was observed between AP and non-AP musicians). Eight right-handed (three males) professional musicians without AP were chosen from a pool of 16 musicians that had taken part in previous experiments [Schulze et al., 2011a,b] to match the AP group for age and gender. Within the non-AP group mean age was 24.8 years (SD = 2.00; range: 22–27). Non-AP musicians had started their musical training at an average age of 6.38 years (SD = 0.92 years). Two participants, who were not native speakers of German (one male, both spoke German fluently), were included in the group of non-AP musicians. Eight non-AP musicians and six AP musicians were right-handed according to the Edinburgh Handedness Inventory [Oldfield, 1971]; mean lateralization quotient was 95% for non-AP musicians, and 56% for AP musicians, with no significant difference between the two groups (Mann-Whitney test: $P = 0.33$). An independent-samples t-test showed no difference between AP and non-AP musicians for the age of musical commencement ($t_{14} = 0.98$, $P = 0.34$). The study was approved by the local ethics committee of the University of Leipzig, and conducted in accordance with the Declaration of Helsinki.

AP Testing

AP was confirmed using two established AP tests [Keenan et al., 2001; Schulze et al., 2009] in which participants had to name sine wave tones (duration of each tone = 500 ms). The first AP test consisted of 13 tones, ranging from F#4 (370 Hz) to F#5 (740 Hz). Each tone was presented four times, resulting in 52 sine wave tones. The second AP test consisted of 12 sine wave tones which were presented four times, resulting in 48 tones to name. The tones in this AP test ranged from C2 (65 Hz) to G5 (1568 Hz). In accordance with previous studies [Keenan et al., 2001; Loui et al., 2010; Miyazaki, 1988; Takeuchi and Hulse, 1993; Ward and Burns, 1982; Wynn, 1992, 1993] answers within one semitone of the presented pitch were regarded as correct answers, that is, the chance level of correct responses was 25%. Participants were instructed to respond as fast and as accurately as possible. Both AP and non-AP musicians performed the AP test. As shown in Figure 1, AP musicians performed significantly better than chance ($P > 0.001$) in both AP tests. They gave 92% (SEM = 2.2%) in the first and 82% (SEM = 5.6%) correct responses in the second AP test. Non-AP musicians performed at chance level in both test ($P > 0.45$ for both). AP musicians responded on average within 1416 ms (SEM = 453 ms) in the first, and 2024 ms (SEM = 902 ms) in the second test (reaction times were measured with regard to the offset of the tone that had to be named).

To investigate the ability of participants to name the tones in more detail, they were given an AP questionnaire. Participants indicated in this AP questionnaire whether they could name some tones played by their instrument. Although non-AP musicians performed at chance in the AP test, six out of the eight non-AP musicians tested indicated that they can name some tones played by their instrument.

Stimuli

To ensure that activation differences were not due to different physical properties of the stimuli, we presented tones and syllables simultaneously. Each auditory stimulus comprised a spoken syllable and a simultaneously presented sine wave tone (see Fig. 2), matched for loudness. The frequencies of the sine wave tones corresponded to

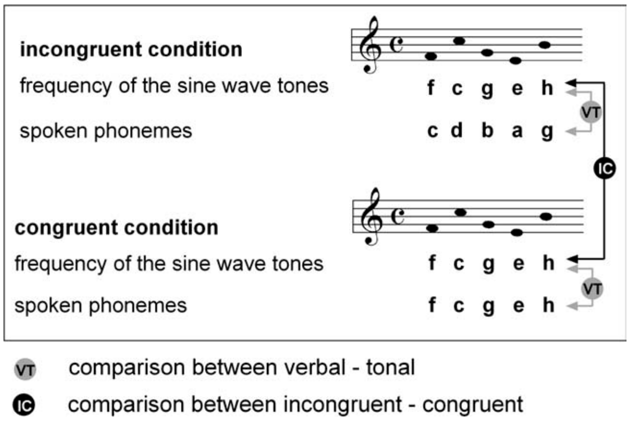

Figure 2.

Experimental conditions. Verbal and tonal condition: Each auditory stimulus comprised of a spoken syllable and a simultaneously presented sine wave tone. The frequencies of the sine wave tones corresponded to the frequencies of the tones of the Western chromatic scale. The syllables were the spoken German names of the scale tones (e.g., gis [g-sharp], c, etc.). Congruent and incongruent condition: In 50% of the sequences the frequency of the sine wave tones corresponded to the simultaneously presented syllables (congruent sequences) and in the remaining 50% of sequences, the frequency of the sine wave tones did not correspond to the syllables (incongruent sequences). Congruent and incongruent conditions were only compared during the tonal condition.

the frequencies of the tones of the Western chromatic scale (based on A4 = 440 Hz), and ranged from 261 Hz to 523 Hz (C4–C5; one octave), resulting in 13 different tones of 12 pitch classes. The syllables were spoken by a professional male speaker and were the German names of the scale tones (e.g., gis [g-sharp], c, etc.), resulting in 12 different syllables (tones with a frequency of 261 Hz and 523 Hz are both referred to as c). In each experimental trial, five such stimuli were presented in a sequence. Each stimulus had duration of 400 ms, with periods of 150 ms of silence between them, resulting in a sequence length of 2600 ms.

Procedure

There were two tasks, a verbal and a tonal task (see Fig. 2 for illustration). In both tasks, participants listened to sequences of five auditory stimuli (the sample stimuli sequence). They were instructed to internally rehearse the syllables during the verbal task or the tones during the tonal task for 4200 up to 6200 ms. At the end of each trial a test stimulus was presented, consisting of the simultaneous presentation of one syllable and one sine wave tone, and participants had to indicate by a button press, whether the syllable (in the verbal task), or the sine wave tone (in the tonal task), had already been presented during the sample stimuli sequence.

Auditory Stroop Task

The syllables were the names of the scale tones. In 50% of the sequences the frequency of the sine wave tones corresponded to the simultaneously presented syllables (congruent sequences) and in the remaining 50% of sequences, the frequency of the sine wave tones did not correspond to the syllables (incongruent sequences). Congruent and incongruent conditions were only compared during the tones condition (see Fig. 2 for a visualization of examples for congruent and incongruent sequences). To ensure that, as in the congruent condition, there would be no difference between the verbal and tonal "contour" in the incongruent sequences, the verbal and tonal contours were matched (e.g., verbal sequence (F G A D B) and tonal sequence (A B C# F# D)).

During a pink noise control condition, participants did not perform a verbal or tonal WM task: Pink noise was presented instead of the sample stimuli sequence and the test stimulus to control for the auditory perceptual input. Additionally, participants pressed a predefined button after the end of the sequence to account for the motor response.

The sequences were presented pseudorandomly in a blocked design, and participants started either with a verbal or tonal block (counterbalanced across participants and groups) for a total of 10 blocks. Each block consisted of 16 experimental sequences (±1 sequence), resulting in 160 experimental sequences, 80 presented during the verbal task, and 80 presented during the tonal task. At the beginning of each block a visual cue indicated whether the WM task for the next block was the verbal or tonal task. In each block, three pink noise control sequences (±1 sequence) were presented additionally, resulting in a total of 30 pink noise control sequences. The blocks and sequences used were identical for the tone and verbal WM task. Participants were repeatedly instructed not to sing or hum aloud during the scanning session.

Behavioral Data Analysis

Performance data (as percentage of correct responses) were analyzed with the SPSS v.16 statistical software package (SPSS, Chicago, IL). If the data were normally distributed (Shapiro-Wilk test) ANOVAs and *t*-tests were used to analyze within-participant and between-participant differences, respectively. If the data were not normally distributed non-parametric test equivalents were used (Wilcoxon tests and Mann–Whitney *U* tests).

Scanning Paradigm

The scanning paradigm was a modified version of the sparse temporal sampling technique [Hall et al., 1999], that is, auditory stimulation was presented in the absence of the scanner noise. Two scans per trial were acquired. This allowed us to scan the hemodynamic response associated with (a) the perception/encoding (first scan) and (b) the rehearsal (second scan) of the sequences. Five different onsets of the auditory sequence relative to the first scan were used to allow optimal sampling of the hemodynamic response associated with the encoding (scans occurred 0, 500, 1000, 1500, or 2000 ms after the auditory presentation). The rehearsal time differed in length accordingly (rehearsal times: 4200, 4700, 5200, 5700, or 6200 ms). The first scan was associated with perceptual/encoding processes, and the second scan captured the BOLD response associated with the WM rehearsal process. This scanning paradigm, which had been used by our group in previous experiments [Schulze et al., 2011a,b], enabled us to analyze data of the first scan (encoding/perception) and second scan (WM rehearsal) separately.

Data Acquisition

The experiment was carried out on a 3T scanner (Siemens TRIO, Erlangen). Before each functional session, an anatomical reference data set was acquired for each participant, which was standardized to the Talairach stereotactic space [Talairach and Tournoux, 1988]. A bunched gradient-echo EPI sequence was used with a TE of 30 ms, a flip angle of 90° and a TR of 6600 ms and an acquisition bandwidth of 100 kHz. 24 axial slices were acquired rapidly within approximately 1600 ms, so that no scanning occurred during the rest of the TR. The matrix dimensions were 64×64 with a field of view (FOV) of 192 mm, resulting in a voxel size of 3×3, slice thickness of 4 mm, and an interslice gap of 1 mm.

Data Analysis

Pre-processing, statistical analysis and visualization of the fMRI data were performed with the software package LIPSIA [Lohmann et al., 2001]. An offline motion correction was performed on the functional images, using the Siemens motion correction protocol. Thereafter, functional slices were aligned to a 3D stereotactic coordinate reference system. The registration parameters were acquired to achieve an optimal match between the functional slices and the individual 3D reference dataset, which was standardized to the Talairach stereotactic space [Talairach and Tournoux, 1988]. In addition to this linear transformation, a non-linear registration was performed between the anatomical 3D datasets of the group [Thirion, 1998], and the resulting deformation fields were applied to the associated functional datasets. In the last step of pre-processing, the data were smoothed with a Gaussian filter of 8-mm full width at half maximum. The statistical evaluation was based on the general linear model [Friston et al., 1998]. Contrast images were generated by computing differences between the parameter estimates (e.g., between the tonal and the verbal condition). To restrict the statistical analysis to relevant voxels inside the brain, a mask was applied to eliminate data from outside the brain and in the ventricular system. In addition, contrast images showing differences between AP and non-AP musicians for the verbal–tonal perception were restricted to the left hemisphere according to our hypothesis [see Introduction, and Hirata et al., 1999; Itoh et al., 2005; Loui et al., 2010; Oechslin, Imfeld et al., 2009; Schulze et al., 2009; Wilson et al., 2009].

Contrast images were entered into a second-level random effects analysis. One-sample *t*-tests were performed to evaluate whether observed differences were significantly different from zero. To protect against false-positive activation, the results were corrected for multiple comparisons by the use of cluster-size and cluster-value thresholds obtained by Monte Carlo simulations with a significance level of $P < 0.05$ [Lohmann et al., 2008]. Clusters were obtained with a voxel-wise threshold of $P < 0.005$ and an extent threshold of 10 voxels [Forman et al., 1995].

The activation differences between verbal and tonal perception; verbal and tonal rehearsal; incongruent and congruent perception; and incongruent and congruent rehearsal were compared between AP and non-AP groups using interaction contrasts (group x condition). Significant activation differences observed in these interaction contrasts were then further investigated with within-group comparisons (verbal vs. tonal perception, verbal vs. tonal rehearsal, incongruent vs. congruent perception, and incongruent vs. congruent rehearsal were compared in AP and non-AP musicians separately) and with between-group comparisons (congruent perception, congruent rehearsal, incongruent perception, incongruent rehearsal, verbal perception, verbal rehearsal, tonal perception, and tonal rehearsal were compared between AP and non-AP musicians). Therefore, within- and between-group analyses were reported in regions showing a significant (group × condition) interaction.

RESULTS

Performance data for both groups is visualized in Figure 3. Performance of AP musicians in the tonal task was significantly better compared to the non-AP musicians ($t_{14} = 2.64$, $P = 0.02$). In the verbal task, performance did not differ between groups ($Z = 0.21$; $P = 0.83$). Significant differences between groups were observed for the congruent condition ($Z = 2.7$; $P = 0.007$), but not for the incongruent condition ($Z = 1.1$; $P = 0.268$; Bonferroni corrected alpha for multiple comparisons: $P < 0.025$). Within-participant differences between the congruent and the incongruent condition were neither observed for the AP musicians ($Z = 1.28$; $P = 0.202$) nor for the non-AP musicians ($Z = 0.49$; $P = 0.622$).

An ANOVA with within-participant factor Material (congruent/incongruent) and between-participant factor AP (AP and non-AP) was used to analyse reaction times. This analysis showed no main effect of Material ($F_{1,14} = 0.055$, $P = 0.82$) or AP ($F_{1,14} = 0.11$, $P = 0.75$), and no Material*AP Interaction ($F_{1,14} = 2.40$, $P = 0.14$).

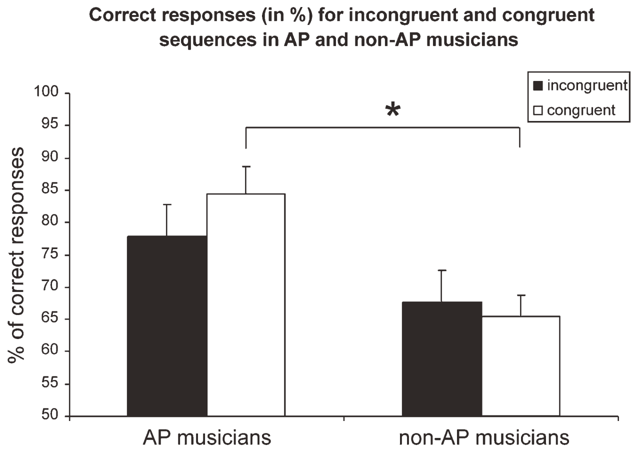

Figure 3.
WM performance (in % correct responses) for AP and non-AP musicians presented as a function of sequence type (congruent/incongruent)—error bars indicate the standard error of mean (SEM).

fMRI Data

Significant activation differences observed in the interaction contrasts were further investigated with within-group and with between-group comparisons (see Materials and Methods). Activations are listed in Table I and are shown in Figure 4 (a complete list of activation differences between and within groups for all conditions is provided in the Supporting Information Table I).

Verbal Versus Tonal Perception and Rehearsal

The statistical analysis for the contrast showing differences between AP and non-AP musicians for the verbal–tonal perception was restricted to the left hemisphere (see Methods). AP musicians showed stronger activation of the left middle occipital gyrus (BA 19) during verbal compared to tonal perception. In non-AP musicians an increased BOLD response was observed in the left anterior STG/STS (BA 22/38) and in the left lingual gyrus (BA 18) during verbal compared to tonal perception (see Fig. 4). Neither for AP musicians nor for non-AP musicians, we observed stronger activation for tonal perception when compared to verbal perception. Furthermore, for both groups, no functional differences between the verbal and the tonal condition were observed during rehearsal.

Incongruent Versus Congruent Perception and Rehearsal

The superior frontal gyrus (SFG; BA 10) bilaterally, the anterior and middle portion of the left STG/STS (BA 22), and the right postcentral gyrus (BA 43) exhibited stronger activation during incongruent compared to congruent perception in AP musicians (see Fig. 4).

No structure showed an increased BOLD response during congruent compared to incongruent perception in AP musicians.

In contrast to AP musicians, who showed stronger activation only during incongruent compared to congruent perception (but not vice versa), non-AP musicians exhibited stronger activations only for congruent compared to incongruent perception.

The left supramarginal gyrus/STG (BA 40/41), the right lingual/fusiform gyrus (BA 19/37), the cuneus (BA 18/19) and the cerebellum bilaterally showed increased activation for the perception of the congruent compared to the incongruent sequences in non-AP musicians.

For both groups, no functional differences were observed comparing congruent and incongruent rehearsal.

Tonal Perception and WM Between AP and Non-AP Musicians

The comparison of tonal perception and WM between AP and non-AP musicians was of interest in the present study because a stronger activation of the left STG/STS was expected [Schulze et al., 2009], but no differences were observed. The complete list of activation differences between and within groups for all conditions is provided in the Supporting Information Table I.

DISCUSSION

Behavioral Data

A better performance for the congruent, but not the incongruent sequences was observed for the AP musicians compared to the non-AP musicians. Our behavioral results therefore indicate that the verbal labelling of the tones in AP musicians is at least partly automatic, even if this is disadvantageous. Thus, these results corroborate findings by Miyazaki [2004] and Itoh et al. [2005]. Overall, AP musicians showed a better performance for tone sequences (incongruent and congruent together) than non-AP musicians, a result which is consistent with the literature [Siegel, 1974; Zatorre et al., 1998]. Furthermore, our data did not show a difference between musicians with and without AP for the verbal sequences, indicating that AP musicians do not simply have a better WM capacity.

fMRI Data

As hypothesized, we observed a stronger BOLD response in the left STG/STS during verbal perception compared to tonal perception in non-AP musicians, but not AP musicians. This result is in accordance with a study [Schulze et al., 2009] that showed a stronger

TABLE I. Within- and between-group (AP and non-AP musicians) activation differences (verbal and tonal perception/rehearsal; incongruent and congruent perception/rehearsal)

		Left hemisphere						Right hemisphere					
		Coordinates *x*, *y*, *z*	*z*-Value	Coordinates *x*, *y*, *z*	*z*-Value	Coordinates *x*, *y*, *z*	*z*-Value	Coordinates *x*, *y*, *z*	*z*-Value	Coordinates *x*, *y*, *z*	*z*-Value	Coordinates *x*, *y*, *z*	*z*-Value
Structure	BA	Interaction contrast group (AP vs. non-AP) × condition		AP		Non-AP		Interaction contrast group (AP vs. non-AP) × condition		AP		Non-AP	
Verbal > tonal perception													
PMd	6	−32, 6, 54	3.84										
Anterior STG/STS	22/38	−50, 3, −6	−3.11			−55, 3, −6	4.01						
Middle occipital gyrus	19	−35, −87, 21	3.47	−35, −87, 21	3.58								
Lingual gyrus	18	−14, −96, −6	−3.00			−17, −99, −3	3.44						
Verbal > tonal rehearsal													
n.s.													
Incongruent > congruent perception													
SFG	9/10	−23, 60, 15	3.15	−23, 60, 12	3.48			7, 57, 27	3.50	10, 57, 24	3.58		
STG/STS	22	−59, −15, 3	3.10	−56, −15, 0	3.82								
Supramarginal gyrus/STG	40/41	−44, −21, 18	3.22			−44, −21, 15	−3.01	43, −18, 18	3.81	43, −15, 21	3.59		
Postcentral gyrus	43							58, −9, 18	4.28	58, −12, 18	4.47		
Lingual/fusiform gyrus	19/37	−32, −57, 0	3.53					40, −45, −9	3.34			37, −48, −12	−3.71
Middle occipital gyrus	19							37, −66, 6	3.30				
Cuneus	18/19	−14, −93, 9	3.31			−14, −96, 12	−2.97	10, −90, 12	3.37			5, −90, 12	−3.36
Cerebellum		−20, −60, −21	4.10			−25, −57, −24	−4.78	22, −54, −42	3.50			22, −54, −45	−4.28
		−5, −63, −3	2.85			−5, −60, −3	−3.09						
		−26, −54, −39	3.50			−26, −54, −36	−3.53						
Incongruent > congruent rehearsal													
n.s.													

In "AP" or "non-AP" columns: positive *z*-values = verbal > tonal or incongruent > congruent; negative *z*-values = tonal > verbal or congruent > incongruent. All activations are corrected for multiple comparisons with a significance level of $P < 0.05$. Clusters were obtained using a voxel-wise threshold of $P < 0.005$ and an extent threshold of 10 voxels. The statistical analysis for the contrast showing differences between AP and non-AP musicians for the verbal>tonal perception was restricted to the left hemisphere (see Methods).

BA, Brodmann area; PMd, dorsolateral premotor cortex; SFG, superior frontal gyrus; STG, superior temporal gyrus; STS, superior temporal sulcus.

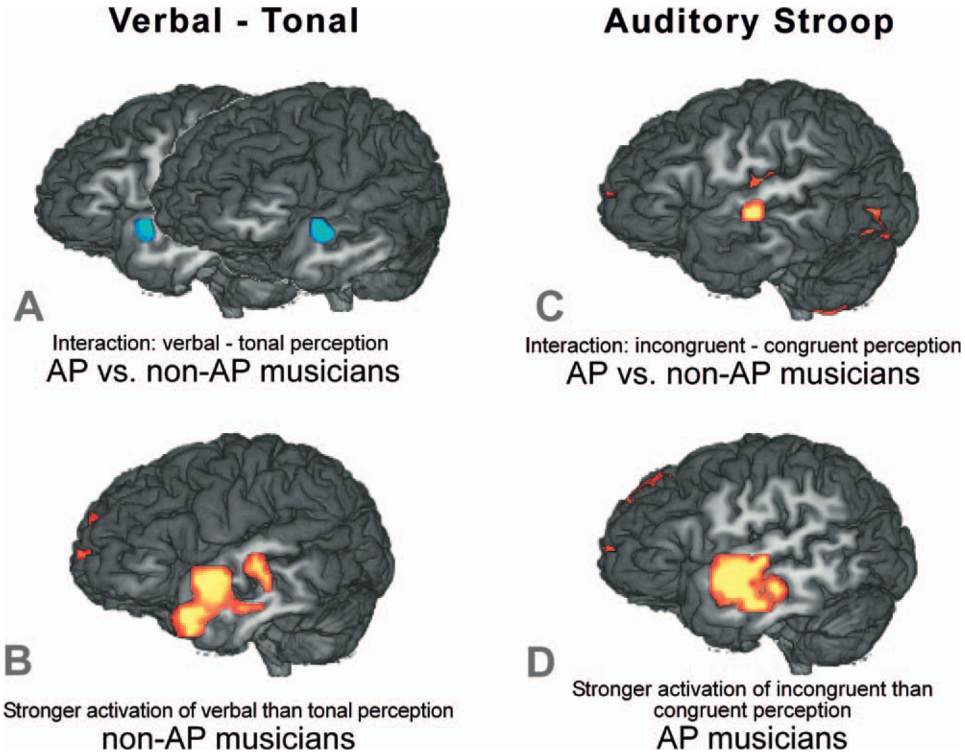

Figure 4.

Significant differences are shown with blue (**A**) or red/yellow (**B**–**D**) color-coding using a cluster threshold of $P < 0.05$, corrected. A – Interaction contrast: difference verbal—tonal was compared between AP and non-AP musicians. B – Stronger activation of left STG/STS during verbal compared to tonal perception in non-AP musicians. C – Interaction contrast: difference incongruent – congruent perception has been compared between AP and non-AP musicians. D – Stronger activation of left STG/STS during incongruent compared to congruent perception in AP musicians.

involvement of the left STS in AP musicians during the perception of tone stimuli, compared to non-AP musicians. This finding is taken here to reflect a more categorical perception in AP musicians based on the following considerations. Humans show a stronger activation of the left anterior and mid-portion of the STG/STS during the perception of verbal material compared to tones [Binder et al., 2000; Liebenthal et al., 2005]. This structure appears to be important in phonemic perception and categorical perception. A more posterior activation of the left STG/STS exhibited stronger activation after a training in which participants learned to perceive stimuli as intelligible speech, compared to a pre-training condition in which the same stimuli could not be understood [Desai et al., 2008; Leech et al., 2009; Mottonen et al., 2006]. Furthermore, Scott and colleagues [e.g., Scott et al., 2000, 2006] demonstrated that the perception of intelligible speech, contrasted against the perception of noise-vocoded speech (that contained some phonetic information), elicited activations in the left anterior STS. Results by Dick et al. [2010] suggest that the left posterior STS region is involved in the categorization of auditory information that is relevant for behavior. This conclusion is based on an increased activation of the left posterior STS in violinists (compared to non-violinists) during listening to music, a region that is typically associated with the analysis of speech.

Because the neural correlates underlying verbal and tonal perception have, to our knowledge, never been compared between AP and non-AP musicians before, our study reports for the first time that this well-known verbal–tonal difference in the left STG/STS is not observed in AP possessors. Thus, our findings indicate that verbal and tonal perception strongly overlap in the left STG/STS in AP musicians, but not in non-AP musicians. This corroborates and extends findings by previous studies: Schulze et al. [2009] reported a stronger activation of the left STS in AP musicians compared to non-AP musicians during tonal perception. However, the activation observed in the present study is more anterior than the activation reported in the previous study [Schulze et al., 2009]. This discrepancy is presumably due to two aspects. First, all tones in the present experiment—in contrast to Schulze et al. [2009]—corresponded to the frequencies of the tones of the Western musical scale and second, the reported contrasts. While Schulze et al. [2009] compared only tonal perception between AP and non-AP musicians, the present study included a verbal condition. In contrast to Schulze et al. [2009], no difference in the left STS was observed in the present study when comparing tonal perception between AP and non-AP musicians. This might be due to the smaller group size in the present study. In another study speed and accuracy during a pitch naming task in AP musicians was associated with activation in the left posterior STG [Wilson et al., 2009] and the authors suggested that this finding points to a neurobiological basis of an AP template in the temporal lobe. These functional differences [Schulze et al., 2009; Wilson et al., 2009] were supported by anatomical differences between musicians with and without AP: Loui et al. [2010] showed that individuals with AP possess enhanced white matter connectivity between STG and MTG, a neural correlate that might underlie the modified perception in AP musicians. Furthermore, whereas Oechslin, Meyer et al. [2009] observed that the ability of AP is associated with different functional responses to verbal material in the temporal lobe, we did not observe differences in the temporal lobe between both groups during the verbal task.

To summarize, the observed overlap between verbal and tonal perception in the left STG/STS associated with AP suggests a categorical perception of tones in AP musicians, comparable to verbal perception. Thus, the findings reported here support the assumption that AP already influences perception [Zatorre, 2003].

The second main result is that incongruent sequences elicited an increased BOLD response in the left anterior STG/STS in AP musicians, but not in non-AP musicians, in a similar region that was more strongly involved during verbal compared to tonal perception in non-AP musicians but not in AP musician. This indicates that the anterior STG/STS is involved in detecting the incongruencies between tones and tone-names. As already discussed above, the middle portion of the STS is involved in phonemic processing [e.g., Binder et al., 2000; Liebenthal et al., 2005; Scott et al., 2000, 2006], categorical perception [Liebenthal et al., 2005], and processing of intelligible verbal material [e.g., Scott et al., 2000, 2006].

Itoh et al. [2005] conducted an EEG study to investigate the auditory stroop effect in AP and non-AP musicians. They observed an ‘AP negativity’ (maximal over left posterior STG regions, with a peak latency of 150 ms), which was elicited during both listening and pitch-naming conditions. Notably, this effect was modulated by stimulus incongruency, and the amplitude of this effect was significantly smaller for incongruent compared to congruent stimuli. Our results together with previous EEG studies [Hirata et al., 1999; Itoh et al., 2005; Wu et al., 2008] corroborate the notion that AP affects tone processing as early as during perception. Furthermore, the present fMRI study revealed that the left STG/STS in AP musicians might underlie the perception of incongruency during an auditory stroop task.

In addition, AP musicians, but not non-AP musicians, showed an increased activation of the superior frontal gyrus (SFG; BA 10), also termed frontopolar cortex, during the perception of incongruent sequences compared to congruent sequences. The frontopolar cortex is known to be involved in a multitude of higher-level cognitive tasks [for an overview see Koechlin and Hyafil, 2007], for example in (i) decision making, especially during more unstructured situations, (ii) the exploration and selection of all possible options and switching back and forth between alternatives, and (iii) solving several internal tasks and integrating their results. Thus, our data indicate that for the AP musicians the incongruent condition increased the cognitive demand in terms of attention, working memory and multitasking, presumably due to the internal discrepancy of the stimuli [for an overview see Gilbert et al., 2006]. The functional difference, that is, a stronger involvement of the frontopolar cortex during the perception of incongruent compared to congruent sequences, together with the behavioral result of a tendency towards a better performance for congruent compared to incongruent sequences, suggest that the verbal labelling of tones is at least partly automatic in AP musicians, even if this is disadvantageous.

Interestingly, even though sine wave tones were used [which are more difficult to name than tones played by an instrument; see Takeuchi and Hulse, 1993] and although we controlled for the contour between tones and tone names in the incongruent condition (see Materials and Methods), a difference in brain activation for non-AP musicians between the incongruent and congruent condition was observed: The left supramarginal gyrus/STG, the right lingual/fusiform gyrus, and bilaterally the cuneus and the cerebellum showed increased activation during the perception of the congruent compared to the incongruent sequences.

These functional differences in non-AP musicians were surprising because (i) non-AP musicians did not show a performance difference between congruent and incongruent sequences and (ii) non-AP musicians’ performed at chance for the AP tests. However, it has been proposed that there are latent forms of AP [for an overview see Vitouch, 2003], which provides a potential explanation for the observed activation difference between the perception of congruent and incongruent sequences in non-AP musicians. This interpretation was corroborated by the answers of the non-AP musicians given for the AP questionnaires: Six out of the eight non-AP musicians tested indicated that they can name some tones played by their instrument (such as the tones of the open strings of a string instrument).

The supramarginal gyrus and the cerebellum, which interestingly showed increased activation during the congruent compared to the incongruent sequences in non-AP musicians, are part of a network underlying verbal and tonal WM [Gaab et al., 2003; Gaab and Schlaug, 2003; Hickok et al., 2003; Koelsch et al., 2009; Schulze et al., 2011a]. Our results indicate that non-AP musicians relied more strongly on these classical WM regions during the congruent compared to the incongruent condition. Activation in the occipital cortex, which again was more strongly activated in non-AP musicians during the congruent condition compared to the incongruent condition, has been observed to be involved in the perception of tones and melodies [Belin et al., 1999; Platel et al., 1997; Satoh et al., 2001], and when musicians imagined to play [Meister et al., 2004] or actually played their instrument [Parsons et al., 2005]. This suggests, together with our results, that some form of visual imagery might have played a role during perceiving congruent, compared to incongruent, sequences in non-AP musicians. Because this study was mainly designed to investigate congruent and incongruent perception and rehearsal in AP musicians, future studies are needed to investigate how congruency influences perception and rehearsal of pitch information in musicians without absolute pitch.

It should be noted that the differences observed between AP and non-AP musicians cannot be due to the fact that German was not the native language for more of the AP musicians because: (a) the participants were studying music in Germany and where fluent German speakers, (b) we did not observe a behavioral difference in WM for verbal material between both groups, and (c) no functional difference in the left STG/STS was observed during verbal perception between AP and non-AP musicians.

It has been suggested that there are two main auditory pathways that both project from the auditory cortex [Hickok and Poeppel, 2007; Rauschecker, 1998; Rauschecker and Scott, 2009; Rauschecker and Tian, 2000]: A dorsal pathway projects into the posterior parietal and frontal cortex and is thought to be responsible for spatial processing [Rauschecker and Scott, 2009] and sensorimotor integration [Hickok and Poeppel, 2007]. A ventral "what" pathway projects into the anterior temporal cortex and is involved in the identification of auditory objects [Hickok and Poeppel, 2007; Rauschecker and Scott, 2009]. Our finding of an absent functional difference between verbal and tonal perception and the stronger activation of the left STG/STS during incongruent compared to congruent sequences in AP musicians suggests that the left anterior STG/STS (i.e., part of the "what" pathway) is in AP musicians also involved in perceiving (and identifying) tones.

It is assumed that the ability of AP does not only influence the perception of tones, but also WM for tones. Several studies [Hantz et al., 1992; Klein et al., 1984; Wayman et al., 1992] reported a lack, or a reduction, of the P300-ERP in AP musicians in a pitch memory task. Because the P300 is an electrophysiological index for WM processes [Klein et al., 1984], this was taken to reflect that AP musicians update their WM less frequently (compared to non-AP musicians) due to internal tone templates [Zatorre et al., 1998]. Schulze et al. [2009] also observed a stronger activation of the right SPL/IPS, possibly indicating a stronger involvement of WM related areas in non-AP musicians compared to non-AP musicians during a WM task for tones. Therefore, we expected differences for tonal WM between AP and non-AP musicians in the present study. However, when comparing tonal or verbal WM between AP and non-AP musicians, no differences were observed. Perhaps the functional differences observed during tonal perception in our study also contributed to a superior performance of AP musicians during the tonal WM task; this issue remains to be specified.

SUMMARY AND CONCLUSIONS

The data reveal that, in the left STG/STS, neural resources for verbal and tonal perception overlap more strongly in AP musicians than in non-AP musicians. Likewise, the left anterior STG/STS is involved in the detection of incongruent stimuli during an auditory stroop task in AP musicians, but not in non-AP musicians. This is taken to reflect that perceiving and identifying tones involves the "what" pathway of the auditory system (of which the anterior STG/STS is part) in AP musicians. Our results also indicate that first, AP is associated with the categorical perception of tones; second, the left STG/STS is activated in AP musicians only for the detection of verbal-tonal incongruencies in the auditory stroop task; and finally, verbal labelling of tones in AP musicians seems to be automatic. Overall, a unique feature of AP appears to be the similarity between verbal and tonal perception.

ACKNOWLEDGMENTS

This work was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, grant KO 2266/4-1 awarded to S.K.) and the Max Planck Society.

REFERENCES

Baddeley AD (2003): Working memory: Looking back and looking forward. Nat Rev Neurosci 4:829–839.

Belin P, Zatorre RJ, Hoge R, Evans AC, Pike B (1999): Event-related fMRI of the auditory cortex. Neuroimage 10:417–429.

Bermudez P, Lerch JP, Evans AC, Zatorre RJ (2009): Neuroanatomical correlates of musicianship as revealed by cortical thickness and voxel-based morphometry. Cereb Cortex 19:1583–1596.

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

Bermudez P, Zatorre RJ (2009): The absolute pitch mind continues to reveal itself. J Biol 8:75.

Binder JR, Frost JA, Hammeke TA, Bellgowan PS, Springer JA, Kaufman JN, Possing ET (2000): Human temporal lobe activation by speech and nonspeech sounds. Cereb Cortex 10: 512–528.

Dehaene-Lambertz G, Pallier C, Serniclaes W, Sprenger-Charolles L, Jobert A, Dehaene S (2005): Neural correlates of switching from auditory to speech perception. Neuroimage 24:21–33.

Desai R, Liebenthal E, Waldron E, Binder JR (2008): Left posterior temporal regions are sensitive to auditory categorization. J Cogn Neurosci 20:1174–1188.

Dick F, Lee HL, Nusbaum H, Price CJ (2010): Auditory-motor expertise alters "speech selectivity" in professional musicians and actors. Cereb Cortex doi:10.1093/cercor/bhq166.

Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC (1995): Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): Use of a cluster-size threshold. Magn Reson Med 33:636–647.

Friston KJ, Fletcher P, Josephs O, Holmes A, Rugg MD, Turner R (1998): Event-related fMRI: Characterizing differential responses. Neuroimage 7:30–40.

Gaab N, Gaser C, Zaehle T, Jancke L, Schlaug G (2003): Functional anatomy of pitch memory-an fMRI study with sparse temporal sampling. Neuroimage 19:1417–1426.

Gaab N, Schlaug G (2003): The effect of musicianship on pitch memory in performance matched groups. Neuroreport 14: 2291–2295.

Gilbert SJ, Spengler S, Simons JS, Steele JD, Lawrie SM, Frith CD, Burgess PW (2006): Functional specialization within rostral prefrontal cortex (Area 10): A meta-analysis. J Cogn Neurosci 18:932–948.

Hall DA, Haggard MP, Akeroyd MA, Palmer AR, Summerfield AQ, Elliott MR, Gurney EM, Bowtell RW (1999): "Sparse" temporal sampling in auditory fMRI. Hum Brain Mapp 7:213–223.

Hantz EC, Crummer GC, Wayman JW, Walton JP, Frisina RD (1992): Effects of musical training and absolute pitch on the neural processing of melodic intervals: A P3 event-related potential study. Music Perception 10:25–42.

Hickok G, Buchsbaum B, Humphries C, Muftuler T (2003): Auditory-motor interaction revealed by fMRI: Speech, music, and working memory in area Spt. J Cogn Neurosci 15:673–682.

Hickok G, Poeppel D (2007): The cortical organization of speech processing. Nat Rev Neurosci 8:393–402.

Hirata Y, Kuriki S, Pantev C (1999): Musicians with absolute pitch show distinct neural activities in the auditory cortex. Neuroreport 10:999–1002.

Hirose H, Kubota M, Kimura I, Ohsawa M, Yumoto M, Sakakihara Y (2002): People with absolute pitch process tones with producing P300. Neurosci Lett 330:247–250.

Itoh K, Suwazono S, Arao H, Miyazaki K, Nakada T (2005): Electrophysiological correlates of absolute pitch and relative pitch. Cereb Cortex 15:760–769.

Jancke L, Wustenberg T, Schulze K, Heinze HJ (2002): Asymmetric hemodynamic responses of the human auditory cortex to monaural and binaural stimulation. Hear Res 170:166–178.

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

Klein M, Coles MGH, Donchin E (1984): People with absolute pitch process tones without producing a P300. Science 223: 1306–1309.

Koechlin E, Hyafil A (2007): Anterior prefrontal function and the limits of human decision-making. Science 318:594–598.

Koelsch S, Schulze K, Sammler D, Fritz T, Muller K, Gruber O. (2009).Functional architecture of verbal and tonal working memory: An FMRI study. Hum Brain Mapp 30:859–873.

Leech R, Holt LL, Devlin JT, Dick F (2009): Expertise with artificial nonspeech sounds recruits speech-sensitive cortical regions. J Neurosci 29:5234–5239.

Liebenthal E, Binder JR, Spitzer SM, Possing ET, Medler DA (2005): Neural substrates of phonemic perception. Cereb Cortex 15:1621–1631.

Lohmann G, Muller K, Bosch V, Mentzel H, Hessler S, Chen L, Zysset S, von Cramon DY (2001): LIPSIA—a new software system for the evaluation of functional magnetic resonance images of the human brain. Comput Med Imaging Graph 25:449–457.

Lohmann G, Neumann J, Mueller K, Lepsien J, Turner R (2008): The multiple comparison problem in fMRI—a new method based on anatomical priors. In: Proceedings of the First Workshop on Analysis of Functional Medical Images. In 11th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI) (New York City).

Loui P, Li HC, Hohmann A, Schlaug G (2010): Enhanced cortical connectivity in absolute pitch musicians: A model for local hyperconnectivity. J Cogn Neurosci 23:1015–1026.

Luders E, Gaser C, Jancke L, Schlaug G (2004): A voxel-based approach to gray matter asymmetries. Neuroimage 22:656–664.

Luo H, Husain FT, Horwitz B, Poeppel D (2005): Discrimination and categorization of speech and non-speech sounds in an MEG delayed-match-to-sample study. Neuroimage 28:59–71.

Meister IG, Krings T, Foltys H, Boroojerdi B, Muller M, Topper R, Thron A (2004): Playing piano in the mind-an fMRI study on music imagery and performance in pianists. Brain Res Cogn Brain Res 19:219–228.

Miyazaki K.1988.Musical pitch identification by absolute pitch possessors. Percept Psychophys 44:501–512.

Miyazaki K (2004): The Auditory Stroop Interference and the Irrelevant Speech/Pitch Effect: Absolute-Pitch Listeners Can't Suppress Pitch Labeling. http://citeseerx.ist.psu.edu/viewdoc/summary?doi10.1.1.117.3584

Mottonen R, Calvert GA, Jaaskelainen IP, Matthews PM, Thesen T, Tuomainen J, Sams M (2006): Perceiving identical sounds as speech or non-speech modulates activity in the left posterior superior temporal sulcus. Neuroimage 30:563–569.

Oechslin MS, Imfeld A, Loenneker T, Meyer M, Jancke L (2009): The plasticity of the superior longitudinal fasciculus as a function of musical expertise: A diffusion tensor imaging study. Front Hum Neurosci 3:76.

Oechslin MS, Meyer M, Jancke L (2009): Absolute pitch—Functional evidence of speech-relevant auditory acuity. Cereb Cortex 20:447–455.

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.

Oldfield R C (1971): The assessment and analysis of handedness: The Edinburgh inventory. Neuropsychologia 9:97–113.

Parsons LM, Sergent J, Hodges DA, Fox PT (2005): The brain basis of piano performance. Neuropsychologia 43:199–215.

Platel H, Price C, Baron JC, Wise R, Lambert J, Frackowiak RS, Lechevalier B, Eustache F (1997): The structural components of music perception. A functional anatomical study. Brain 120: 229–243.

Rakowski A (1993): Categorical perception in absolute pitch. Arch Acoust 18:515–523.

Rauschecker JP (1998): Cortical processing of complex sounds. Curr Opin Neurobiol 8:516–521.

Rauschecker JP, Scott SK (2009): Maps and streams in the auditory cortex: Nonhuman primates illuminate human speech processing. Nat Neurosci 12:718–724.

Rauschecker JP, Tian B (2000): Mechanisms and streams for processing of "what" and "where" in auditory cortex. Proc Natl Acad Sci USA 97:11800–11806.

Satoh M, Takeda K, Nagata K, Hatazawa J, Kuzuhara S (2001): Activated brain regions in musicians during an ensemble: A PET study. Brain Res Cogn Brain Res 12:101–108.

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

Schulze K, Gaab N, Schlaug G (2009): Perceiving pitch absolutely: Comparing absolute and relative pitch possessors in a pitch memory task. BMC Neurosci 10:106.

Schulze K, Zysset S, Mueller K, Friederici AD, Koelsch S (2011a): Neuroarchitecture of verbal and tonal working memory in nonmusicians and musicians. Hum Brain Mapp 32:771–783.

Schulze K, Mueller K, Koelsch S (2011b): Neural correlates of strategy use during auditory working memory in musicians and non-musicians. Eur J Neurosci 33:189–196.

Scott SK, Blank CC, Rosen S, Wise RJ (2000): Identification of a pathway for intelligible speech in the left temporal lobe. Brain 123:2400–2406.

Scott SK, Rosen S, Lang H, Wise RJ (2006): Neural correlates of intelligibility in speech investigated with noise vocoded speech—a positron emission tomography study. J Acoust Soc Am 120:1075–1083.

Siegel JA (1974): Sensory and verbal coding strategies in subjects with absolute pitch. J Exp Psychol 103:37–44.

Specht K, Rimol LM, Reul J, Hugdahl K (2005): "Soundmorphing": A new approach to studying speech perception in humans. Neurosci Lett 384:60–65.

Takeuchi AH, Hulse SH (1993): Absolute pitch. Psychol Bull 113:345–361.

Talairach P, Tournoux J (1988): A Stereotactic Coplanar Atlas of the Human Brain. Stuttgart: Thieme.

Thirion JP (1998): Image matching as a diffusion process: An analogy with Maxwell's demons. Med Image Anal 2:243–260.

Vitouch O (2003): Absolutist models of absolute pitch are absolutely misleading. Music Percep 21:111–118.

Ward W, Burns E (1982): Absolute pitch. In: Deutsch D, editor. The Psychology of Music. New York: Academic Press. pp431–451.

Wayman JW, Frisina RD, Walton JP, Hantz EC, Crummer GC (1992): Effects of musical training and absolute pitch ability on event-related activity in response to sine tones. J Acoust Soc Am 91:3527–3531.

Wilson SJ, Lusher D, Wan CY, Dudgeon P, Reutens DC (2009): The neurocognitive components of pitch processing: Insights from absolute pitch. Cereb Cortex 19:724–732.

Wu C, Kirk IJ, Hamm JP, Lim VK (2008): The neural networks involved in pitch labeling of absolute pitch musicians. Neuroreport 19:851–854.

Wynn VT (1992): Absolute pitch revisited. Br J Psychol 83129–131.

Wynn VT (1993): Accuracy and consistency of absolute pitch. Perception 22:113–121.

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, Beckett C (1989): Multiple coding strategies in the retention of musical tones by possessors of absolute pitch. Mem Cognit 17:582–589.

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.