Cerebral Cortex February 2010;20:447–455
doi:10.1093/cercor/bhp113
Advance Access publication July 10, 2009

Absolute Pitch—Functional Evidence of Speech-Relevant Auditory Acuity

Mathias S. Oechslin, Martin Meyer and Lutz Jäncke

Department of Neuropsychology, University of Zurich, CH-8050 Zurich, Switzerland

Absolute pitch (AP) has been shown to be associated with morphological changes and neurophysiological adaptations in the planum temporale, a cortical area involved in higher-order auditory and speech perception processes. The direct link between speech processing and AP has hitherto not been addressed. We provide first evidence that AP compared with relative pitch (RP) ability is associated with significantly different hemodynamic responses to complex speech sounds. By systematically varying the lexical and/or prosodic information of speech stimuli, we demonstrated consistent activation differences in AP musicians compared with RP musicians and nonmusicians. These differences relate to stronger activations in the posterior part of the middle temporal gyrus and weaker activations in the anterior mid-part of the superior temporal gyrus. Furthermore, this pattern is considerably modulated by the auditory acuity of AP. Our results suggest that the neural underpinnings of pitch processing expertise exercise a strong influence on propositional speech perception (sentence meaning).

Keywords: absolute pitch, fMRI, musical expertise, speech processing

Introduction

"Absolute pitch" (AP) is a very rare phenomenon among professional musicians, enabling them to identify tones without the aid of any reference tone. In terms of cognitive music psychology, AP could be characterized as the ability to distinguish and identify one salient quality (the pitch chroma) from a number of other perceptual attributes (Levitin and Zatorre 2003), which constitute the conflated unity of complex sounds. The prevalence among professional musicians differs between cultures: Prevalence rates in Japan have been reported up to 50% (Miyazaki 1988; Gregersen et al. 1999) compared with estimates of 1–20% for professional Western musicians (Vitouch 2003). It has been suggested that this effect is associated with the Suzuki method (Gregersen et al. 2001), which is a widespread pedagogical music approach in Japan originally intended for violin training The Suzuki music education emphasizes learning music by ear over reading musical notation and preferably begins with formal lessons early in life between the ages of 3 and 5 years. One fundamental reasoning in favor of this education points to a parallelism between natural speech acquisition and purely auditory based musical training because the former also goes without any visual cues and is exclusively based on auditory feature learning (Kuhl 2003). Hence, akin to language acquisition where a child learns to understand and to produce spoken language before learning to read the Suzuki music education strives for acquisition of musical skills based on pure auditory sensation and production of music before learning to read music. Interestingly, memorization of musical pieces without referring to a notation is an important key issue of this approach which particularly puts emphasis on auditory features, namely discrimination and representation of pitch and timbre. This makes children who started musical training due to the Suzuki method early in their life ideal candidates when it comes to studying the influence of auditory focused musical training on the development of the auditory system. Moreover, it has been shown that the influence of an acquired tonal language like Mandarin can have a considerable impact on the development of AP: The prevalence of AP was far greater among the Chinese than the US students for each age level of musical training onset (Deutsch et al. 2006). Thus these authors suggested that a tonal language enables infants to associate pitches with verbal labels during the critical period in which features of their native language are acquired.

Whether the extraordinary ability of AP is genetically determined or develops under the influence of environmental factors has attracted much debate (Vitouch 2003; Zatorre 2003; Levitin and Rogers 2005). Irrespective of the much disputed role of the former (Athos et al. 2007; Drayna 2007), there is considerable evidence for the substantial impact of early musical training on the development of AP (Baharloo et al. 1998; Russo et al. 2003; Miyazaki and Ogawa 2006). Presently there is a broad consensus that automatic language acquisition more likely yields a native-speaker proficiency when it occurs before a critical age (Lenneberg 1967; Johnson and Newport 1989; Newport 1990; Kuhl 2000; Sakai 2005). On the other hand, with respect to the development of AP, various studies suggest, that the acquisition of AP is strongly related to early musical exposure before the age of 6–7 (Baharloo et al. 1998, 2000; Costa-Giomi et al. 2001; Gregersen et al. 2001). The extent to which language acquisition and AP ability develop in parallel suggests that AP ability might be regarded as a model of cortical plasticity for deliberate practice and musicianship. Thus it is assumed that the proficiency of musicians with AP should result in pertinent characteristics of auditory related cortical areas. Consistent with this notion are various structural observations of morphological changes in the cortical region of the planum temporale (PT) in musicians with AP (Schlaug et al. 1995; Keenan et al. 2001; Luders et al. 2004; Wilson et al. 2008) and Heschl's gyrus (Schneider et al. 2005) in musicians with relative pitch (RP). These studies show that AP and professional musicianship in general leads to marked cortical gray matter alterations, mainly in the form of left-hemispheric asymmetries in speech-relevant areas. The pivotal role of the PT in auditory processing has been supported by a review article, in which the PT is taken to be a *computational hub* (Griffiths and Warren 2002) that is involved in processing different types of complex acoustic signals like environmental sounds, speech and music. Functional studies to date have revealed hemodynamic and electrophysiological lateralization effects in musicians with AP during processing of musical stimuli (Hirata et al. 1999; Ohnishi et al. 2001; Schneider et al. 2005; Gaab et al. 2006; Wu et al. 2008). By comparing nonmusicians and musicians with RP, several electrophysiological studies demonstrated a higher level of pitch decoding performance in language as a function of musical expertise (Marques et al. 1981; Schön et al. 2004; Itoh et al. 2005; Magne et al. 2006). There are however different levels of musical expertise, including superior AP, the functional importance of which is still unclear for speech processing. Because basic auditory processing is crucial for both speech and music perception, the relationship between musical expertise and speech processing needs to be examined. Prosodic variations, that is natural pitch modulations in spoken sentences, share many acoustic features with tone transitions in musical melodies, which are mostly characterized as slow temporal variations of spectral units that span over several segments (Meyer et al. 2002). The processing of delexicalized speech, effectively pure speech prosody, leads bilaterally to a significantly reduced neural activity of the posterior superior temporal gyrus (STG), PT and the planum polare (Meyer et al. 2003, 2004). Besides these slow temporal variations—taken to be *suprasegmental* information in the presented stimuli—there are also fast temporal changes, signaling important information in speech and music: rapid spectrotemporal signal changes constitute the *segmental* information of speech (Shannon et al. 1995; Hickok and Poeppel 2007). In particular, important phonemic cues are indicated by these rapid signal changes. Without the ability to perceive them one would probably be unable to discriminate, for example, between the words *peer* and *beer*. Precise phonetic processing and full dynamic lexical access therefore more efficiently contribute to a comprehensive semantic understanding (Poeppel et al. 2008).

Regarding the neural basis of language comprehension, a linguistically based model of spoken language comprehension discriminates between *segmental* and *suprasegmental* information of speech (Friederici and Alter 2004). Particularly in terms of the *dynamic dual pathway model* the authors argue, that *segmental* information (phonemes, syntactic and lexial–semantic elements) are primarily processed in a left hemisphere temporo-frontal pathway whereas *suprasegmental* information (sentence level prosody) is processed in a right hemispheric temporo-frontal pathway. Moreover, the authors imply dynamic interactions between the hemispheres, due to a disentangling of prosodic and semantic information during auditory sentence comprehension. With respect to the neural processing of *segmental* speech, it has been shown (Meyer et al. 2004) that the left hemisphere STG and superior temporal sulcus (STS) activations are most strongly driven by *segmental* information processing irrespective of whether the presented speech stimuli comprise prosodic pitch variations (that is *suprasegmental* information) or not. This finding goes in line with the *dual stream model* proposed by Hickok and Poeppel (2007). Their *dual stream model of cortical organization of speech processing* assumes a *dorsal stream*, which is mainly involved in speech production, connecting left-hemispheric posterior supratemporal regions with inferior frontal areas. Complementarily they claim a *ventral stream*, which in principle is bilaterally represented. The *ventral stream* is thought to be responsible for a mediation of spectrotemporal (STG) and phonological (STS) analyses with lexical units located in inferior temporal regions. Moreover, a proposed lexical interface (middle temporal gyrus [MTG], inferior temporal sulcus) subserves these processes by gating and collating basic auditory and lexical memory information.

This proposed link between basal auditory and higher-order speech information processing leads to the main hypothesis of this paper assuming a link between musical expertise (especially for AP musicians) and higher-order (lexical and/or prosodic) speech information processing.

A recently published diffusion tensor imaging study (Glasser and Rilling 2008) focused on 2 distinct seeding regions (STG and STS/MTG) in the left hemisphere in order to track the superior longitudinal fasciculus (SLF)—the main pathway in association with speech perception and production. The authors overlaid activation findings from other studies using either lexical, phonemic or prosodic language stimuli. They showed that activations based on lexical speech stimuli corresponded to the MTG seeding region of interest (ROI), whereas phonemic processing was associated with the STG seeding ROI. Thus, the authors demonstrated a left lateralized functional association of lexical speech processing by taking into account the morphological architecture of the SLF. Furthermore, we have been able to demonstrate that local alterations of diffusion parameters among the SLF are associated with key regions (like MTG and inferior frontal gyrus [IFG]) by means of higher-order language processing, and modulated by different levels of musical expertise (Oechslin et al. forthcoming).

Applying brain imaging methods musicians with AP have not been studied so far with respect to speech processing. To date all imaging studies published have used musical stimuli during functional MRI. Exemplary Ohnishi and colleagues (Ohnishi et al. 2001) observed enhanced responses in the left PT while AP musicians listened to melodies, whereas other studies have shown that the right auditory cortex is preferentially activated when nonmusicians process music (Tervaniemi et al. 2001; Janata et al. 2002; Overy et al. 2004).

Based on these findings showing enhanced responses in the left PT and adjacently located perisylvian brain regions in AP musicians to musical stimuli we assume that there is an increased proficiency of AP individuals also in language processing. In particular, we posit a left-sided lateralization also in language comprehension as a function of musical expertise irrespective of linguistic domain (syntax, semantics, phonology).

This idea is supported by several studies which reported anatomical and functional alterations in left-sided perisylvian brain areas of AP musicians (Schlaug et al. 1995; Steinmetz 1996; Keenan et al. 2001; Luders et al. 2004; Wilson et al. 2008). It is also conceivable that AP musicians use their augmented pitch memory (Gaab et al. 2006) ability to more efficiently identify linguistically relevant pitch information than do RP and nonmusicians (NM). If this is indeed the case, AP musicians might show less activation in left-sided perisylvian brain areas when processing linguistic speech stimuli.

In order to elucidate the possible link between the acuity of AP and speech perception we designed the present functional magnetic resonance imaging (fMRI) experiment in which meaning and intonation in spoken language were systematically varied. We were specifically interested in whether AP musicians demonstrate different cortical activation patterns in association with lexical as compared with prosodic speech information. Because AP musicians demonstrate particular morphological and functional alterations in the left PT region, we reasoned that left-sided perisylvian and adjacent extrasylvian areas (STS, MTG) would be differently involved in higher-order speech processing. Therefore, we anticipated that these differences would occur within the ventral pathway as delineated by Hickok and Poeppel (Hickok and Poeppel 2007).

Materials and Methods

Subjects and AP Test

Fifteen professional musicians with AP (8 females/7 males; mean age = 24 years, SD = 4.2; mean practice years = 18.4, SD = 2.9; mean age of practice begin = 5.7, SD = 2,2), fifteen professional musicians with RP (8 females/7 males, mean age = 25.3 years, SD = 2.8; mean practice years = 16.6, SD = 3.8; mean age of practice begin = 8.7, SD = 3) and fifteen NM without any musical expertise (NM: 8 females/7 males, mean age = 25.7, SD = 5.4) participated in this study. NM were selected on the basis that they had no musical practice for at least fifteen years. None of the subjects reported any hearing impairments. All participants were tested for their handedness with the Annett Handedness Inventory (Annett 1967). All of them had normal structural scans and did not suffer from any neurological disorders. We evaluated AP among all participating professional musicians with an in-house test: participants heard 108 pure sine wave tones, presented in pseudorandomized order, which ranged from A3 (tuning: A4 = 440 Hz) to A5, while each tone was presented 3-fold. The accuracy was evaluated by counting correct answers—the semitone errors were taken as incorrect to increase the discriminatory power. Furthermore, the participants were not asked to identify the adjacent octaves of the presented tones, as for AP it is a most notable prerequisite to identify the correct chroma. Accurate detection of octaves is quite a difficult task, which is hardly possible even for musicians with AP.

Each tone of the AP test had a duration of 1 s; the interstimulus interval (ISI) of 4 s was filled with brown noise. Subjects had to write down the tonal label immediately after they heard the accordant tone (i.e., while hearing the 4 s of brown noise). The whole test unit and its components were created with Adobe Audition 1.5. The AP test was performed with a Dell Laptop Latitude 300x and presented via Sennheiser HD-25-1 headphones.

The Experimental Procedure

The 4 conditions (Fig. 1), which encompass the manipulation of spoken German phrases are characterized as follows: *normal speech* (yielding proper propositional speech), *delexicalized speech* (representing pure speech melody/prosody or pitch contour), *flattened speech* (representing pure lexical and syntax information—comprising sentence meaning and lacking dynamic pitch contour) and *flattened-delexicalized speech* (combined application of the prior 2 manipulations, lacking both sentence meaning and dynamic pitch contour). Delexicalizement of speech signals leads to a masking of lexical and syntactic information as a result of the PURR-filtering procedure (Sonntag and Portele 1998). This kind of manipulation produces speech stimuli containing only prosodic speech parameters such as intonation, duration, amplitude envelope and the second and third formants. The procedure to generate *flattened speech* is based on a readjustment of the pitch contour F_0, in which all natural pitch variations are kept constant on the level of 200 Hz. All stimuli were normalized on the same moderate amplitude level. These 4 conditions are conceived as the expression of 2 dimensions representing prominent speech inherent characteristics, namely *segmental* and *suprasegmental* information. *Delexicalized speech* and *flattened speech* represent the first and the second dimension of our experimental design and are each defined by 2 levels (*suprasegmental* on/off and *segmental* on/off). The third dimension is defined by *expertise*—the between-subject variable—which determines 3 levels of musical expertise: AP, RP, and musically untrained subjects. Given this experimental design our fMRI data analysis was performed by means of a 3-way ANOVA with repeated measurements *segmental* × *suprasegmental* × *expertise*). This ANOVA reveals

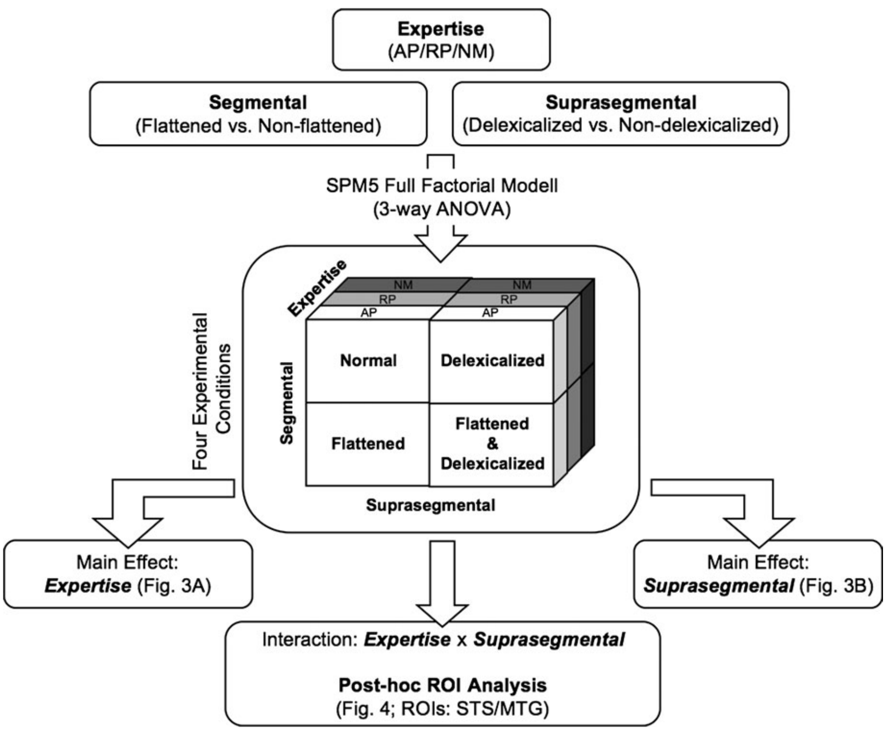

Figure 1. In this figure the methodical framework is depicted. The 3 factors *expertise*, *suprasegmental* and *segmental* leads us to an orthogonal design that has been calculated by using a full factorial design (3-way ANOVA), provided by SPM5: *expertise* (AP/RP/NM) × *segmental* (flattened vs. nonflattened) × *suprasegmental* (delexicalized vs. nondelexicalized). The significant interaction *expertise* × *suprasegmental* has been further analyzed by applying a post hoc ROI analysis comparing delexicalized versus nondelexicalized conditions.

cortical activation clusters that represent significant main effects and interactions respectively. To assure the participants' attention, all subjects were instructed to judge whether a sentence contains suprasegmental information (i.e., prosody) or not, and to respond via a response box after each trial. Each condition comprised 40 sentences. Stimuli of the 4 conditions being presented in pseudorandomized order, evenly distributed in 4 separate runs (each 10.6 min). Each sentence had a duration of about 5 s followed by an ISI of 11 s, resulting in a trial length of 16 s—an adequate time window to model the blood oxygenation level–dependent (BOLD) response. The sentences started in a jittered order to preserve the variance within the BOLD signal (onset-times for sentences: 1, 500, 1000, 1500, 2000 ms).

Data Acquisition and Analysis

During the scanning session the participants were instructed to keep their eyes open and to focus a fixation cross. Binaural auditory stimulation was presented by a digital playback system including a high frequency shielded transducer system. This acoustic transmission system includes a piezoelectric loudspeaker enabling the transmission of strong sound pressure levels (105 dB) with excellent attenuation characteristics (Jäncke et al. 2001). T_2^*-weighted echo planar imaging (EPI) was acquired on a 3.0 tesla GE magnet resonance scanner (imaging parameters: echo time = 32 ms, repetition time = 2 sec, flip-angle = 70 deg., FOV = 22 cm, slice thickness = 3.4 mm, voxel size = 3.4 mm × 3.4 mm, slices per volume = 32, volumes = 302). The data analysis was performed with the parametric mapping software SPM5 (http://www.fil.ion.ucl.ac.uk/spm/). The preprocessing consisted of spatial realignment, normalization to a standard EPI template and a smoothing procedure with a 6-mm Gaussian kernel. Due to the experimental design, the analysis was proceeded in an event-related manner; therefore the standardized canonical HRF was applied to model the BOLD response. For further group level analysis we specified the SPM5 factorial design built up by 3 independent variables resulting in a 2 × 2 × 3 ANOVA: *segmental* (2 levels: un-/flattened sentences), *suprasegmental* (2 levels: un-/delexicalized sentences) and *expertise* (3 levels: AP/RP/NM). The reported main effects and interactions are all proceeded on the $P < 0.001$ level (unc.) with an extended cluster threshold of $k = 5$ voxels. Furthermore, to elucidate hemispheric asymmetries during speech processing as a function of musical expertise we performed a post hoc ROI-analysis regarding the interaction *segmental* × *expertise*. The software *marsbar* (http://marsbar.sourceforge.net/) was used to define 7-mm sphere ROIs bilaterally at maximal *F*-values reflecting the 2 predefined left-hemispheric clusters (Fig. 4*A*: ROI 1, STS, [−54, −37, 6]; ROI 2, MTG, [−51, −39, −6]). Mean BETA values were read out by in-house programmed MATLAB (http://www.mathworks.com/) scripts and further analyzed by a general linear model with repeated measures and *t*-tests (SPSS, http://www.spss.com/).

Results

Forty-five healthy volunteers participated in our study. They were grouped according to 3 distinctive levels of musical expertise: AP possessors, RP possessors, and nonmusicians (NM) without any musical expertise as controls. The professional musicians (AP/RP) performed an in-house designed AP test. Using a behavioral AP performance index (Fig. 2, AP test score), 2 distinct experimental groups were formed. The data show a clear distinction between the 2 groups, whereas AP accuracy is heterogeneously distributed within the groups (AP: $n = 15$, mean = 82.2%, SD = 16.2; RP: $n = 15$, mean = 6.9%, SD = 4.2). Due to the conservative scoring technique used in this experiment, the performance data indicate that most of the AP musicians have high AP ability. However, the subjects' scores speak against an all-or-none dichotomy regarding the special phenomenon of AP.

To maintain motivation during the fMRI task procedure, subjects were asked to pay attention to the prosodic aspect of the spoken German sentences and to judge each stimulus as to whether it contained prosody or not. Subjects heard 4 different types of acoustic stimuli which were distinguishable along the 2 independent dimensions of *segmental* and *suprasegmental* speech information. Furthermore, the third dimension is defined as *expertise*, enabling (Fig. 1) a 3-way ANOVA (for further specifications see *the experimental procedure* in the methods section). The prosody detection task was not selective and resulted in a ceiling level of accuracy, irrespective of condition and experimental group.

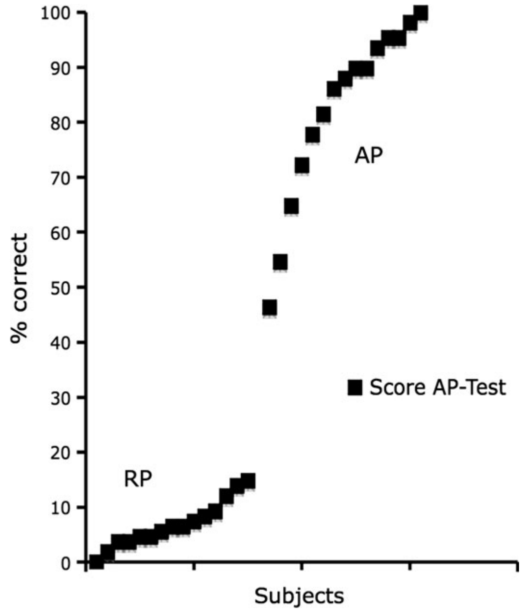

Figure 2. Plotted scores of the AP test (AP [$n = 15$, Average: 82.2%, SD: 16.2] and RP [$n = 15$, Avg.: 6.9%, SD: 4.2]).

All significant clusters representing main effects and interactions based on the performed SPM5 full factorial design (3-way ANOVA) are listed in Table 1. The main effect *expertise* (Fig. 3*A*) is characterized by a bilateral activation of the STG (STG-right; [63, −12, 3], F = 11.79; STG-left; [−57, −9, 3], *F* = 11.32), with the peak of the main effect in the right hemisphere STG. The plotted mean BETA values (Fig. 3*A*) show exactly the same activation pattern in the comparison of the 3 groups over the 4 conditions, the activation in the right STG cluster is considerably enhanced compared with the left STG cluster. This main effect was observed in each condition, the weakest activations were found in musicians with AP.

The significant main effect *suprasegmental* (Fig. 3*B*) reflects predominately left-hemispheric activation in perisylvian regions that constitute the core language network, namely the STS, MTG, IFG, and the inferior temporal gyrus (ITG). The robust main effect *suprasegmental* (activation peak at the left MTG; [−57, −45, 0], *F* = 88.89) can be explained by systematically enhanced brain responses to meaningful sentences. In other words, when comparing the conditions comprising lexical information (*normal* and *flat*) with the conditions lacking lexical information (*delexicalized* and *flattened-delexicalized*), much larger BOLD signals were elicited in the former condition, and here again with a stable distribution across the levels of musical expertise.

The observation of an interaction between lexical information processing and expertise deserves particular consideration: The significant interactions *suprasegmental* × *expertise*

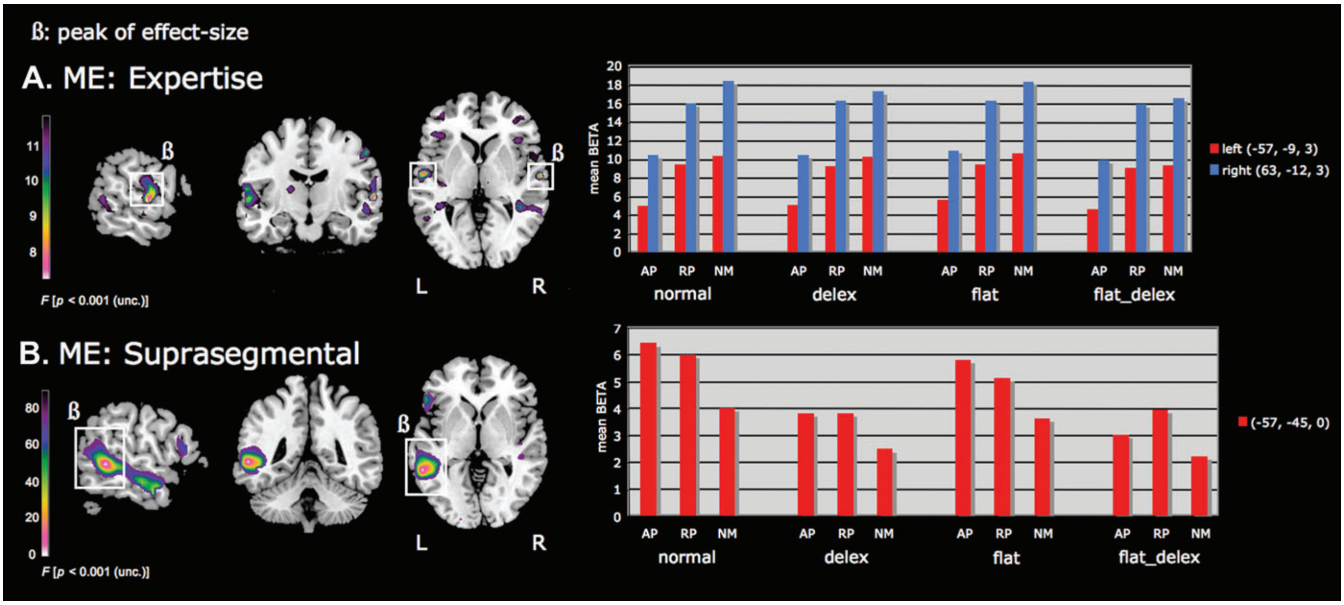

Figure 3. Selected results of the 3-way ANOVA (*segmental* × *suprasegmental* × *expertise*) On the left side, cortical views show the significant results of a full factorial design performed with SPM5: (*A*) the main effect *expertise* (STG, PT) and (*B*) the main effect *suprasegmental* (MTG, STG, ITG). On the right, mean BETA values at the sites of effect peaks (white small boxes) are plotted for all 3 groups of subjects (AP/RP/NM) and against the 4 experimental conditions: *normal speech* (normal), *delexicalized speech* (delex), *flattened speech* (flat) and *flattened-delexicalized speech* (flat_delex).

uncovered a left temporal cluster which is located on the transition strip between the lower bank of the posterior STS and the superior bank of the MTG [(−54, −37, 6), $F = 9.28$]—as this cluster is located considerably inside the sulcus we henceforth use the term STS. Additionally we found a more anterior located cluster on the MTG [(−51, −39, −6), $F = 9.28$]; both clusters are characterized by the same effect size of interaction (Fig. 4*A*). The precise anatomical location was evaluated by applying the *Harvard-Oxford cortical structural atlas* (available at: http://www.cma.mgh.harvard.edu/) and the *Destrieux-Atlas* (Fischl et al. 2004), which has been implemented in *FreeSurfer* software (available at: http://surfer.nmr.mgh.harvard.edu/).

We also conducted a ROI-analysis to more closely examine the relationship between musical expertise and segmental information processing (Figs 1, 4).

First, we defined 2 ROIs based on the interaction peaks of the 2 clusters (Fig. 4*A*: ROI 1, STS, [−54, −37, 6]; ROI 2, MTG, [−51, −39, −6]).

Secondly, we created 2 sphere-ROIs (Bosch 2000) to investigate lateralization effects (Fig. 4*B*).

And thirdly, we conducted four 2-way ANOVAs (*hemisphere* × *expertise*) with repeated measurements based on the mean BETA values for each ROI under each processing condition, that is, nondelexicalized processing (collapsed data of *normal speech* and *flattened speech*) and delexicalized processing (collapsed data of *delexicalized speech* and *flattened-delexicalized speech*) (Fig. 1).

The ANOVA revealed for ROI 1 a main effect for *hemisphere* in the delexicalized conditions ($F_{1,42} = 43.6$, $P < 0.001$); the analysis of ROI 1 obtained in the nondelexicalized conditions revealed a significant main effect for *hemisphere* ($F_{1,42} = 67.2$, $P < 0.001$) and an interaction *hemisphere* × *expertise* ($F_{2,42} = 4.3$, $P < 0.05$). The BETA values for ROI 2 revealed a main effect for *expertise* both in the delexicalized conditions ($F_{2,42} = 3.4$, $P < 0.05$) and in the nondelexicalized conditions ($F_{2,42} = 4.3$, $P < 0.05$). In order to further specify the interaction effects, post hoc tests were conducted (corrected for multiple comparisons) (Fig. 4*C*). The main effect *segmental* did not reveal any suprathreshold cluster ($P < 0.001$ [unc.]).

These findings can be summarized as follows: The main effect *hemisphere* is explained by a strongly left-sided lateralization of activation in the STS. The *hemisphere* × *expertise* interaction relies on the fact that the AP group shows significantly stronger activity in the left than in the right hemisphere during the presentation of segmental speech information. Thus, the STS should be considered an area that supports higher auditory function in AP possessors. However, it should be mentioned that we did not find any interhemispheric difference in the MTG. There was also a main effect for *expertise* in the MTG as shown with post hoc tests revealing higher mean BETA values in musicians than in nonmusicians, whereas there were no significant interhemispheric activation differences. Interestingly, the AP group did not differ in the MTG from the RP group of the musicians. However, the AP musicians showed a highly significant activity enhancement in the left-hemispheric MTG when comparing nondelexicalized with delexicalized categories of stimulus manipulations (Fig. 4*D*).

Discussion

Based on recent studies (Schön et al. 2004; Magne et al. 2006; Wong et al. 2007) one might expect an enhanced sensitivity for pitch contours in musicians during prosodic processing. However, the present study did not reveal a significant main effect when prosody is manipulated (Table 1: main effect [ME] segmental). Notably, as the methodological approaches (event related potentials, brainstem-potentials) and tasks in aforementioned studies clearly differ from our design it is difficult to compare the results, last but not least due to temporal constraints associated with fMRI and the BOLD signal. Unlike the previously mentioned electroencephalography (EEG) studies, the fMRI

Table 1

Significantly activated brain regions broken down for the main effects (ME) and interactions (INT)

Regions	Voxels	F	Coordinates LH			Coordinates RH		
			x	y	z	x	y	z
ME expertise								
STG/PT	14	11.79	—	—	—	63	−12	3
STG/PT	18	11.32	−57	−9	3	—	—	—
RO		8	−60	−9	12	—	—	—
MTG	19	10.89	—	—	—	54	−42	−3
MTG (subgyral)		9.16	—	—	—	42	−42	0
PCG	9	9.58	—	—	—	54	−9	48
ME suprasegmental								
MTG	745	88.89***	−57	−45	0	—	—	—
MTG		61.77***	−54	−15	−15	—	—	—
ITG		56.15***	−57	−6	−24	—	—	—
MTG		54.77***	−51	6	24			
STG		26.24***	−51	−51	18			
IFG	194	46.43***	−54	24	3	—	—	—
IFG		38.91***	−51	18	15	—	—	—
ITG	77	46.14***	—	—	—	57	−9	−21
IFG	32	45.4***	−48	27	−9	—	—	—
SFG	34	28.61**	−9	57	33	—	—	—
SFG		16.7	−12	60	24	—	—	—
PRE	93	20.73	−9	−75	39	—	—	—
PRE		17.83	−18	−66	24	—	—	—
mdFG	39	19.08	−6	39	30	—	—	—
mdFG		14.02	—	—	—	6	39	27
AG	29	18.49	−51	−60	39	—	—	—
PRE	51	17.2	—	—	—	18	−63	21
C		15.4	—	—	—	15	−72	36
IPS	20	15.93	−39	−54	51	—	—	—
MTG	25	14.98	—	—	—	48	−33	−6
mdFG	5	12.35	−3	6	18	—	—	—
ME segmental								
No suprathreshold voxels								
INT suprasegmental × expertise								
MTG	8	9.28	−51	−39	−6	—	—	—
STS	10	9.28	−54	−57	6	—	—	—

Note: This table reports all significant clusters revealed by the SPM full factorial design ($k = 5$, $P < 0.001$ [unc.]), what corresponds to a 3-way ANOVA (*Expertise* × *Segmental* × *Suprasegmental*) Main effects (ME) of *Expertise*, *Suprasegmental*, *Segmental*, and the interaction (INT) *Expertise* × *Suprasegmental* are specified by anatomical labels, cluster size (voxels), the local peak effects (*F* value), and the coordinates of the local peak in the left (LH) and right (RH) hemisphere, respectively. Astersiks ([***] $P < 0.001$, [**] $P < 0.01$) indicate significant clusters due to correction for multiple tests (FWE). PCG, postcentral gyrus; RO, rolandic operculum; IPS, intraparietal sulcus; SFG, superior frontal gyrus; PRE, precuneus; mdFG, medial frontal gyrus; AG, angular gyrus; C, Cuneus.

technique focuses on quite different time frames due to coarser resolution during cortical speech processing. Given the results of these studies, the question may be raised whether AP musicians show an exceptional sensitivity for prosodic processing—however, AP should be considered a phenomenon that differs from standard musical proficiency (which is the subject of investigation in the above cited studies) and may not imply an enhanced sensitivity to prosodic information.

In our statistical analysis the main effect of *expertise* is characterized by bilateral activation on the superior temporal plane with a slight right-hemispheric preponderance (Fig. 2*A*). Musical expertise, irrespective of whether the presented stimuli contained prosodic/lexical information does not account for this finding. The activation pattern is characterized as follows: the higher the musical training—in particular with respect to AP—the lower the activation in the specific region of primary auditory information processing. Accordingly, musical expertise is the main driving factor explaining different activations in the core auditory regions, thus extending recent findings of other research groups (Schneider et al. 2002; Wong et al. 2008).

The main effect of *expertise* possibly indicates more efficient processing by the auditory cortex as a function of musical proficiency. Using EEG, positron emission tomography, and fMRI, the principles of neural efficiency have been discussed within several different contexts such as spatial perception (Vitouch et al. 1997), superior cognitive performance by figural intelligence in chess players (Grabner et al. 2006) and working memory (Grabner et al. 2004). The data provided by these studies imply that higher performance levels are associated with lower cortical activations. Additionally, this relationship has been found (Haier et al. 1992) even as a consequence of "Tetris" learning effects, which were associated with a decrease of local glucose metabolic rates. In general, these authors propose that "brighter" (or more proficient) subjects have to invest less cortical resources to achieve accurate performances. Conversely, cognitively less proficient subjects had to invest more cortical resources to solve the same tasks. From an anatomical point of view, it has been demonstrated, that more gray matter in distinct cortical regions (primary auditory cortex and PT amongst others) is associated with higher IQs; more gray matter also results in less use of energy, when the area is engaged (efficiently) in specific cognitive tasks (Haier et al. 2004).

This argumentation is in line with studies focusing on sensory information processing in the visual cortex (Marcar et al. 2004a). The *standard model* put forward by these authors holds that an increase in the electrical activity and an increase in size of the activated neural population have an opposing influence on the BOLD signal amplitude (Marcar and Loenneker 2004). In a nutshell, this model states that the vascular response is controlled by electrical discharge activity, whereas the oxygen consumption is dependent on the size of the activated neuronal population. Based on an experimental MR setting different checkerboard patterns were presented (flashed vs. reversing), whereas the size of activated neural populations has been manipulated (Marcar et al. 2004b). The results demonstrated that the checkerboard which is associated with a lower neural activity yielded a larger number of activated voxels and a stronger BOLD response.

These results are contradictory to the *Linear Transfer Model* which states that the BOLD contrast signal is directly proportional to the neuronal activity.

In the context of language comprehension, it has been demonstrated that the neural correlates of semantic priming support the neural efficiency hypothesis (Rissman et al. 2003): Semantically related word pairs showed consistently less activation than unrelated pairs—interestingly, with respect to the temporal lobe this activation pattern is restricted to the left STG, and does not affect the MTG. As the authors argue, the perception of a prime word activates a lexical–semantic network that shares common elements with the target word, and, thus, the target can be recognized with enhanced neural efficiency. The proposed relationship between STG and MTG is striking, because the MTG—which doubtlessly is crucial for lexical–semantic processes—does not contribute to the pattern of neural efficiency, drawn by the recruitment of the neural population, which is responsible for primary auditory signal decoding.

Furthermore, in interpreting the main effect of *expertise*, which is characterized by lowest activation of the STG bilaterally in AP and highest activations in NM, it might be useful to recall the characteristic morphological lateralization of Heschl's Gyrus (Schneider et al. 2005) in professional musicians and the PT in professional musicians with AP

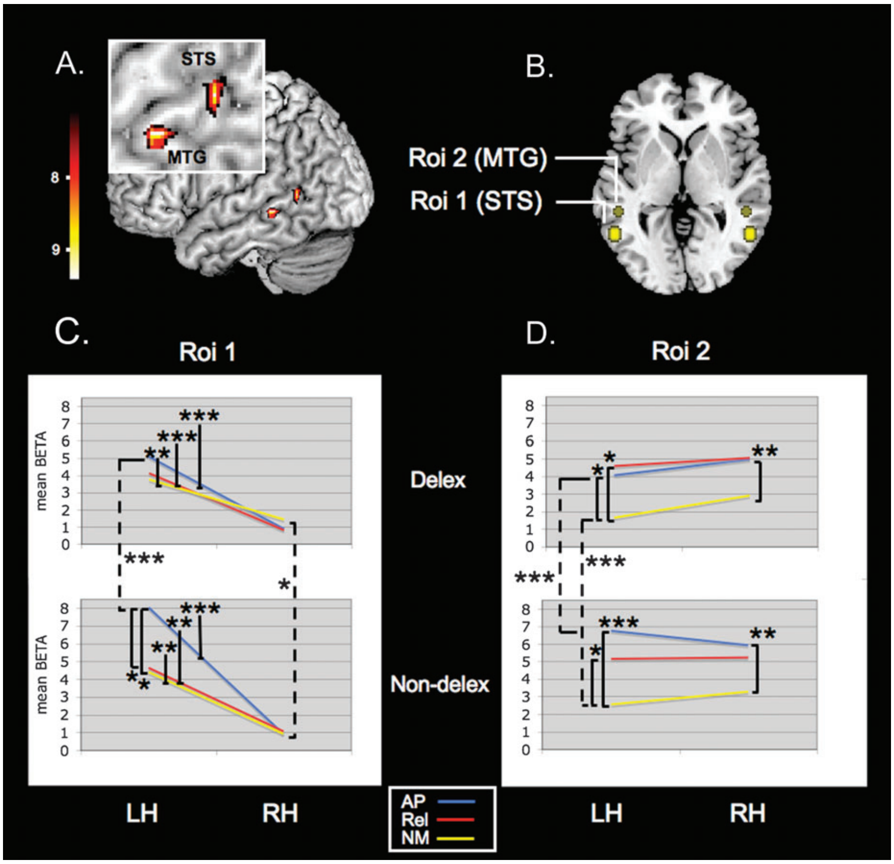

Figure 4. Detailed data of ROI-analysis regarding the significant interaction *suprasegmental* × *expertise* (*A*) STS and MTG interaction cluster with two equivalent left-hemispheric peaks of interaction (ROI 1 [STS]: (−54, −57, 6), $F = 9.28$; ROI 2 (MTG): (−51, −39, −6), $F = 9.28$, $P < 0.001$, $k = 5$), (*B*) two post hoc defined ROIs according to the peaks of interaction at left STS and MTG (left hemisphere) and two corresponding mirror related ROIs (right hemisphere). The left two (C) and right two (D) plots are defined by the separately assigned two clusters of significant interaction (ROI 1: [−54, −57, 6], $F = 9.28$; ROI 2: [−51, −39, −6]). The upper two plots represent the mean BETA values for the collapsed delexicalized conditions (*delexicalized speech* and *flattened-delexicalized speech*) in the left (LH) and the right (RH) hemisphere respectively, and the lower two plots represent the mean BETA values for the collapsed nondelexicalized conditions (*normal speech* and *flattened speech*); asterisks indicate significant levels (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$) as revealed by un-/paired *t*-tests.

(Schlaug et al. 1995; Keenan et al. 2001). Due to an increased size of these structures in the left hemisphere, it seems reasonable to assume that they subserve auditory processing by means of an optimal contribution, characterized by the above-discussed neural efficiency hypothesis. Thus, the revealed main effect of *expertise* can be taken to be the first evidence for neural efficiency in basal auditory processing of language as a function of musical expertise.

When considering the main-effect *suprasegmental*, left temporal brain areas comprising posterior parts of the STG, STS, MTG, and ITG are more strongly activated during processing of lexical and propositional information (Fig. 3*B*). Consistent with this is the finding that the posterior MTG and the lower bank of the posterior STS are involved in lexical and phonetic analyses (Binder et al. 2000; Dick et al. 2007). Essentially, many authors of clinical and nonclinical studies have maintained that the left MTG plays a specific role in lexical and semantic processing (Binder et al. 2000; Dick et al. 2007). By applying *voxel-based lesion-symptom mapping* it has been shown that the posterior MTG is one of the main areas involved in higher-order language processing (Bates et al. 2003; Dronkers et al. 2004). In addition, functional neuroimaging studies support the special role of the posterior MTG in language processing. At least one study that investigated word ambiguity (Rodd et al. 2005) showed elegantly the involvement of the posterior MTG in semantic analyses. Accordingly, there should be a strong link between the auditory cortex and the posterior MTG during lexical information processing. Therefore, both the auditory cortex and the MTG are essential for a proper distinguishing between the words *beer* and *peer*. With respect to this we found a strong interaction between the factors *suprasegmental* and *expertise* located in the left MTG, which is characterized by stronger responses to lexical compared with delexicalized information, with musicians showing the strongest difference. With respect to the main effect *expertise* this activation pattern does not conflict with the above-discussed efficiency hypothesis. As already mentioned, due to the cortical recruitment of lexical-semantic networks it has been demonstrated that primary auditory processing is driven by the principles of neural efficiency, whereas the activity of the MTG shows a different activation pattern which cannot be explained using this line of argumentation. Nevertheless, the MTG provides core evidence for lexical-semantic processing (Rissman et al. 2003).

In general, these results are also in line with the *dual stream model* (Hickok and Poeppel 2007) that postulates a lexical interface located in the posterior part of the left MTG (part of the *ventral stream*). Contrary to strong left-sided activations of the posterior STS due to segmental speech processing (Friederici and Alter 2004), activations of the posterior MTG are specific to musicianship, with musicians (with AP or RP) demonstrating stronger bilateral hemodynamic responses compared with nonmusicians. According to the 2-way ANOVAs (motivated by the findings due to the *expertise* × *suprasegmental* interaction, see Fig. 1), the main effect *expertise* in both delexicalized and nondelexicalized conditions lets us suggest that the MTG might be crucial for higher-level language processing. The post hoc tests show a significant enhancement of effect sizes in this area in musicians compared with nonmusicians (Fig. 4*C*). In addition, the analysis of posterior STS (Fig. 4*C*) revealed that AP musicians show stronger left lateralized activations during processing of segmental information compared with the other 2 groups. This finding is in line with several reports of left-sided enhanced levels of activation in AP musicians during complex auditory tasks (Pantev et al. 1998; Ohnishi et al. 2001; Itoh et al. 2005). Based on these findings, we propose that the auditory acuity of AP is not limited to basal auditory processing (usually conceived in terms of music processing), but extends to a more general notion of acoustic segmentation by fully integrating left-hemispheric speech-relevant networks.

Taken together, our study presents 2 novel findings: First, there is an AP-specific enhancement of the left lateralized activation in the lower bank of the posterior STS for segmental speech processing; second, musicians generally demonstrate stronger bilateral BOLD effects in the posterior MTG in all conditions. In addition, this effect of segmental processing is substantially enhanced in AP musicians compared with the other 2 experimental groups. This novel insight lets us conclude that neurofunctional alterations due to musicianship are not only manifested in exceptional acuity of music processing, but also affect speech processing in the sense that AP represents a comprehensive analytical proficiency for acoustic signal decoding.

Funding

Swiss National Foundation (SNF 46234101) to L.J. and M.O. and (SNF 320000-120661/1) to M.M.; and Research Funding University of Zurich (56234101) to M.O.

Notes

We would like to express our gratitude to the Max Planck Institute for Human Cognitive and Brain Sciences for providing the stimulus material. Moreover, we are indebted to Nicki Langer for assisting in MR data acquisition and to Tino Zaehle for providing support on the data analysis. We also thank 4 anonymous reviewers for their constructive comments on an earlier version of this manuscript. *Conflict of Interest*: None declared.

Address correspondence to Lutz Jäncke, PhD, Department of Neuropsychology, University of Zurich, Binzmühlestrasse 14, CH-8050 Zurich, Switzerland. Email: l.jaencke@psychologie.uzh.ch.

References

Annett M. 1967. The binomial distribution of right, mixed and left handedness. Q J Exp Psychol. 19:327–333.

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.

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.

Bates E, Wilson SM, Saygin AP, Dick F, Sereno MI, Knight RT, Dronkers NF. 2003. Voxel-based lesion-symptom mapping. Nat Neurosci. 6:448–450.

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.

Bosch V. 2000. Statistical analysis of multi-subject fMRI data: assessment of focal activations. J Magn Reson Imaging. 11:61–64.

Costa-Giomi E, Gilmour R, Siddell J, Lefebvre E. 2001. Absolute pitch, early musical instruction, and spatial abilities. Ann N Y Acad Sci. 930:394–396.

Deutsch D, Henthorn T, Marvin E, Xu H. 2006. Absolute pitch among American and Chinese conservatory students: prevalence differences, and evidence for a speech-related critical period. J Acoust Soc Am. 119:719–722.

Dick F, Saygin AP, Galati G, Pitzalis S, Bentrovato S, D'Amico S, Wilson S, Bates E, Pizzamiglio L. 2007. What is involved and what is necessary for complex linguistic and nonlinguistic auditory processing: evidence from functional magnetic resonance imaging and lesion data. J Cogn Neurosci. 19:799–816.

Drayna DT. 2007. Absolute pitch: a special group of ears. Proc Natl Acad Sci USA. 104:14549–14550.

Dronkers NF, Wilkins DP, Van Valin Jr RD, Redfern BB, Jaeger JJ. 2004. Lesion analysis of the brain areas involved in language comprehension. Cognition. 92:145–177.

Fischl B, van der Kouwe A, Destrieux C, Halgren E, Segonne F, Salat DH, Busa E, Seidman LJ, Goldstein J, Kennedy D, et al. 2004. Automatically parcellating the human cerebral cortex. Cereb Cortex. 14:11–22.

Friederici AD, Alter K. 2004. Lateralization of auditory language functions: a dynamic dual pathway model. Brain Lang. 89:267–276.

Gaab N, Schulze K, Ozdemir E, Schlaug G. 2006. Neural correlates of absolute pitch differ between blind and sighted musicians. Neuroreport. 17:1853–1857.

Glasser MF, Rilling JK. 2008. DTI tractography of the human brain's language pathways. Cereb Cortex. 18:2471–2482.

Grabner RH, Fink A, Stipacek A, Neuper C, Neubauer AC. 2004. Intelligence and working memory systems: evidence of neural efficiency in alpha band ERD. Brain Res Cogn Brain Res. 20:212–225.

Grabner RH, Neubauer AC, Stern E. 2006. Superior performance and neural efficiency: the impact of intelligence and expertise. Brain Res Bull. 69:422–439.

Gregersen PK, Kowalsky E, Kohn N, Marvin EW. 1999. Absolute pitch: prevalence, ethnic variation, and estimation of the genetic component. Am J Hum Genet. 65:911–913.

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 Hum Genet. 98:280–282.

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

Haier RJ, Jung RE, Yeo RA, Head K, Alkire MT. 2004. Structural brain variation and general intelligence. Neuroimage. 23:425–433.

Haier RJ, Siegel BV, Tang C, Lennart A, Buchsbaum MS. 1992. Intelligence and changes in regional cerebral glucose metabolic rate following learning. Intelligence. 16:415–426.

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.

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

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

Jäncke L, Gaab N, Wustenberg T, Scheich H, Heinze HJ. 2001. Short-term functional plasticity in the human auditory cortex: an fMRI study. Brain Res Cogn Brain Res. 12:479–485.

Johnson JS, Newport EL. 1989. Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. Cogn Psychol. 21:60–99.

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

Kuhl PK. 2000. A new view of language acquisition. Proc Natl Acad Sci USA. 97:11850–11857.

Kuhl PK. 2003. Human speech and birdsong: communication and the social brain. Proc Natl Acad Sci USA. 100:9645–9646.

Lenneberg E. 1967. Biological foundations of language. New York: Wiley.

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

Levitin DJ, Zatorre RJ. 2003. On the nature of early music training and absolute pitch: a reply to Brown, Sachs, Cammuso, and Folstein. Music Percept. 21:105–110.

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

Magne C, Schoen D, Besson M. 2006. Musician children detect pitch violations in both music and language better than nonmusician children: behavioral and electrophysiological approaches. J Cogn Neurosci. 18:199–211.

Marcar VL, Loenneker T. 2004. The BOLD response: a new look at an old riddle. Neuroreport. 15:1997–2000.

Marcar VL, Strassle A, Girard F, Loenneker T, Martin E. 2004a. When more means less: a paradox BOLD response in human visual cortex. Magn Reson Imaging. 22(4):441–450.

Marcar VL, Strassle AE, Loenneker T, Schwarz U, Martin E. 2004b. The influence of cortical maturation on the BOLD response: an fMRI study of visual cortex in children. Pediatr Res. 56:967–974.

Marques C, Moreno S, Mazzucchi A, Parma M, Cattelani R. 1981. Hemispheric dominance in the perception of tonal sequences in relation to sex, musical competence and handedness. Cortex. 17:291–302.

Meyer M, Alter K, Friederici AD. 2003. Functional MR imaging exposes differential brain responses to syntax and prosody during auditory sentence comprehension. J Neurolinguistics. 16:277–300.

Meyer M, Alter K, Friederici AD, Lohmann G, von Cramon DY. 2002. FMRI reveals brain regions mediating slow prosodic modulations in spoken sentences. Hum Brain Mapp. 17:73–88.

Meyer M, Steinhauer K, Alter K, Friederici AD, von Cramon DY. 2004. Brain activity varies with modulation of dynamic pitch variance in sentence melody. Brain Lang. 89:277–289.

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

Miyazaki K, Ogawa Y. 2006. Learning absolute pitch by children: a cross-sectional study. Music Percept. 24:63–78.

Newport EL. 1990. Maturational constraints of language learning. Cogn Sci. 14:11–28.

Oechslin MS, Imfeld A, Meyer M, Loenneker T, Jäncke L. The plasticity of the superior longitudinal fasciculus as a function of musical expertise: a diffusion tensor imaging study. Forthcoming.

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.

Overy K, Norton AC, Cronin KT, Gaab N, Alsop DC, Winner E, Schlaug G. 2004. Imaging melody and rhythm processing in young children. Neuroreport. 15:1723–1726.

Pantev C, Oostenveld R, Engelien A, Ross B, Roberts LE, Hoke M. 1998. Increased auditory cortical representation in musicians. Nature. 392:811–814.

Poeppel D, Idsardi WJ, van Wassenhove V. 2008. Speech perception at the interface of neurobiology and linguistics. Philos Trans R Soc Lond B Biol Sci. 363:1071–1086.

Rissman J, Eliassen JC, Blumstein SE. 2003. An event-related FMRI investigation of implicit semantic priming. J Cogn Neurosci. 15:1160–1175.

Rodd JM, Davis MH, Johnsrude IS. 2005. The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. Cereb Cortex. 15:1261–1269.

Russo F, Windell D, Cuddy L. 2003. Learning the special note: evidence for a critical period for absolute pitch acquisition. Music Percept. 21(1):119–127.

Sakai KL. 2005. Language acquisition and brain development. Science. 310:815–819.

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

Schneider P, Scherg M, Dosch HG, Specht HJ, Gutschalk A, Rupp A. 2002. Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. Nat Neurosci. 5:688–694.

Schneider P, Sluming V, Roberts N, Scherg M, Goebel R, Specht HJ, Dosch HG, Bleeck S, Stippich C, Rupp A. 2005. Structural and functional asymmetry of lateral Heschl's gyrus reflects pitch perception preference. Nat Neurosci. 8:1241–1247.

Schön D, Magne C, Besson M. 2004. The music of speech: music training facilitates pitch processing in both music and language. Psychophysiology. 41:341–349.

Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M. 1995. Speech recognition with primarily temporal cues. Science. 270: 303–304.

Sonntag GP, Portele T. 1998. PURR—a method for prosody evaluation and investigation. J Comput Speech Lang. 12:437–451.

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

Tervaniemi M, Rytkonen M, Schröger E, Ilmoniemi RJ, Näätänen R. 2001. Superior formation of cortical memory traces for melodic patterns in musicians. Learn Mem. 8:295–300.

Vitouch O. 2003. Absolutist models of absolute pitch are absolutely misleading. Music Percept. 21:111–117.

Vitouch O, Bauer H, Gittler G, Leodolter M, Leodolter U. 1997. Cortical activity of good and poor spatial test performers during spatial and verbal processing studied with slow potential topography. Int J Psychophysiol. 27:183–199.

Wilson SJ, Lusher D, Wan CY, Dudgeon P, Reutens DC. 2008. The neurocognitive components of pitch processing: insights from absolute pitch. Cereb Cortex. 19:724–732.

Wong PC, Skoe E, Russo NM, Dees T, Kraus N. 2007. Musical experience shapes human brainstem encoding of linguistic pitch patterns. Nat Neurosci. 10:420–422.

Wong PCM, Warrier CM, Penhune VB, Roy AK, Sadehh A, Parrish TB, Zatorre RJ. 2008. Volume of left Heschl's Gyrus and linguistic pitch learning. Cereb Cortex. 18:828–836.

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

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.