cortex 59 (2014) 126–137

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Research report

Music listening engages specific cortical regions within the temporal lobes: Differences between musicians and non-musicians

Arafat Angulo-Perkins [a], *William Aubé* [b,c], *Isabelle Peretz* [b,c], *Fernando A. Barrios* [a], *Jorge L. Armony* [b,c,d] *and Luis Concha* [a,b,*]

[a] *Instituto de Neurobiología, Universidad Nacional Autónoma de México. Querétaro, Querétaro, México*
[b] *International Laboratory for Brain, Music and Sound (BRAMS), Montreal, Québec, Canada*
[c] *Department of Psychology, Université de Montréal, Montreal, Québec, Canada*
[d] *Douglas Institute and Department of Psychiatry, McGill University, Montreal, Québec, Canada*

ARTICLE INFO

Article history:
Received 19 July 2013
Reviewed 23 October 2013
Revised 22 February 2014
Accepted 18 July 2014
Action editor Henry Augustus Buchtel
Published online 12 August 2014

Keywords:
Auditory cortex
Music
planum polare
Speech
fMRI

ABSTRACT

Music and speech are two of the most relevant and common sounds in the human environment. Perceiving and processing these two complex acoustical signals rely on a hierarchical functional network distributed throughout several brain regions within and beyond the auditory cortices. Given their similarities, the neural bases for processing these two complex sounds overlap to a certain degree, but particular brain regions may show selectivity for one or the other acoustic category, which we aimed to identify. We examined 53 subjects (28 of them professional musicians) by functional magnetic resonance imaging (fMRI), using a paradigm designed to identify regions showing increased activity in response to different types of musical stimuli, compared to different types of complex sounds, such as speech and non-linguistic vocalizations. We found a region in the anterior portion of the superior temporal gyrus (aSTG) (*planum polare*) that showed preferential activity in response to musical stimuli and was present in all our subjects, regardless of musical training, and invariant across different musical instruments (violin, piano or synthetic piano). Our data show that this cortical region is preferentially involved in processing musical, as compared to other complex sounds, suggesting a functional role as a second-order relay, possibly integrating acoustic characteristics intrinsic to music (e.g., melody extraction). Moreover, we assessed whether musical experience modulates the response of cortical regions involved in music processing and found evidence of functional differences between musicians and non-musicians during music listening. In particular, bilateral activation of the *planum polare* was more prevalent, but not exclusive, in musicians than non-musicians, and activation of the right posterior portion of the superior temporal gyrus (*planum temporale*) differed between groups. Our results provide evidence of functional specialization for

---

[*] *Corresponding author*. Instituto de Neurobiología, Laboratorio C13, Universidad Nacional Autónoma de México, Campus Juriquilla, Boulevard Juriquilla 3001, CP 76230, Querétaro, Querétaro, México.
E-mail address: lconcha@unam.mx (L. Concha).

http://dx.doi.org/10.1016/j.cortex.2014.07.013

music processing in specific regions of the auditory cortex and show domain-specific functional differences possibly correlated with musicianship.

1. Introduction

Through an evolutionary process, the environment has shaped, refined, and modified the neural systems involved in perception. As a result, the human auditory cortex processes not only environmental cues, but also other complex acoustical signals, such as speech and music (Masataka, 2009). Speech is the most relevant sound for modern humans, and this has led many to believe that music is merely its byproduct. This is partially supported by two concepts: 1) music ability is shared with the language faculty (e.g. Patel, 2003); and 2) music is not adaptive (Patel, 2007; Pinker, 1999). Others (e.g., Peretz & Coltheart, 2003), however, have proposed that functional networks specific to either music or speech exist, in order to account for the frequent observation of individuals who, after bilateral lesions to regions of the auditory cortex, can no longer understand speech but can enjoy music and *vice versa* (Peretz, 2006; Wan & Schlaug, 2010).

Two alternate views have emerged in an attempt to address the question of music-specificity of neural networks: The first view, prompted by the shared syntactic integration resource hypothesis for language and musical analysis (Patel, 2003), is based on the syntactic overlap between music and speech processing, and suggests that this phenomenon is mirrored at the neuronal level. Supporting this theory, Abrams et al. (2011) showed data proving an overlap in temporal regions in the processing of these two types of stimuli; however, the authors acknowledged that differences might exist at small scales within these large common areas. The second view suggests a functional specialization for music processing, as can be seen from specific deficits secondary to neurological pathologies (e.g., *amusia*), but also in the strong evidence from different functional paradigms exploring the processing of music and speech (Brown, Martinez, & Parsons, 2006; Rogalsky, Rong, Saberi, & Hickok, 2011; Schmithorst, 2005), and even in pathologies not directly related to music, such as autism (Lai, Pantazatos, Schneider, & Hirsch, 2012).

To further test these discrepant views, we used functional magnetic resonance imaging (fMRI) to compare the brain regions engaged in the processing of each category of stimuli that included acoustic stimulation with a wide range of complex sounds such as human vocal sounds (e.g., with or without linguistic content), and musical sounds of different timbre and rhythms. In particular, our primary goal was to determine whether there are specific temporal regions that preferentially respond to musical stimuli, as compared to other complex acoustic stimuli including speech and nonlinguistic human vocalizations. In order to evaluate the sensitivity to music more widely, other ecologically relevant stimuli were included, such as non-vocal sounds produced by humans, and monkey vocalizations.

Finally, we also assessed whether these putative music-selective (or other) regions are modulated by prior musical training, since previous studies have revealed that musical abilities can modify the distribution of the functional networks and the neuroanatomical characteristics associated with their processing (Herdener et al., 2010; Pantev & Herholz, 2011; Schlaug, Jäncke, Huang, Staiger, & Steinmetz, 1995; Schlaug, Jäncke, Huang, & Steinmetz, 1995). To achieve this goal, our sample included individuals with and without prior formal musical training. This group comparison allowed us to look for differences in music processing based on musical training but also to determine whether intensive music training modulates the activation of areas devoted to speech processing, as observed in other studies of visual and audiomotor expertise (Dick, Lee, Nusbaum, & Price, 2011; Elmer, Hänggi, Meyer, & Jäncke, 2014; Harley et al., 2009).

2. Materials and methods

2.1. *Subjects*

This study included 53 healthy, right-handed volunteers, age 28 ± 8 years (range: 21–55 years; 24 women), all native Spanish speakers. Of these, 25 volunteers were non-musicians (age 29 ± 9 years; range: 21–55 years; 13 women), who had not received extra-curricular music education beyond the mandatory courses in school, and 28 were musicians (age 28 ± 7 years; range: 21–50 years; 11 women), defined as those subjects with at least 3 years of formal studies in music (either instruments or singing) and who were currently involved in musical activities on a daily basis. Groups did not differ in terms of age or gender. All subjects reported normal hearing, which was confirmed during an audio test within the scanner (see below).

All volunteers gave informed consent before the scanning session, and were free of contraindications for MRI scanning. The research protocol had approval from the Ethics Committee of the Institute of Neurobiology at the Universidad Nacional Autónoma de México and was conducted in accordance with the international standards of the Declaration of Helsinki of 1964.

2.2. *Experimental design*

The experiment used a pseudo-randomized block design, with each 10-sec block including 5 stimuli from the same category (Fig. 1). The stimulation protocol had a total duration of 8 min. Stimuli were presented binaurally through MRI-compatible headphones (Nordic NeuroLab, Bergen, Norway) to reduce acoustic interference (~20 dB) generated by the gradients. The volume was adjusted by the volunteers after a short audio test

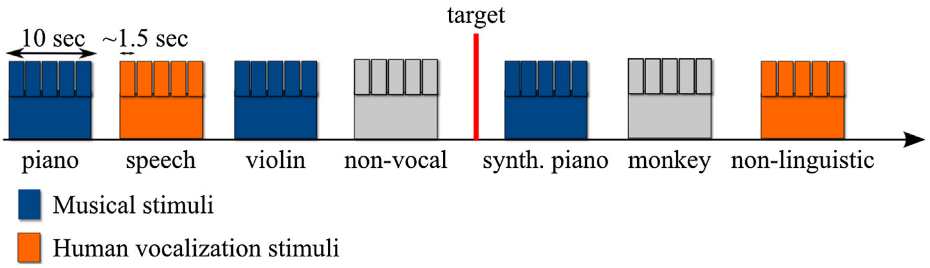

Fig. 1 – Stimulation paradigm and matrix design. Upper panel: Each block includes 5 different stimuli (mean stimulus duration 1.5 ± .2 sec) from the same category. The paradigm included 4 blocks (~10 sec) from each of the 7 categories and had a total duration of 8 min.

in the scanner (volume was deemed comfortable, yet sufficiently high to mask the noise generated by the fMRI acquisitions, ~80 dB).

To ensure subjects were paying attention to the stimuli, they were instructed to press a button with the index finger of their right hand every time they heard a pure tone (500 Hz, 500 msec duration), which was presented 5 times at random during the paradigm. A pure tone was chosen because it is acoustically simple, has a single frequency component, does not vary in amplitude, and is clearly different from the rest and thus easily recognizable by all the subjects. During presentation of the stimuli subjects kept their eyes open staring at a white cross on a black background.

2.3. *Acoustic stimuli*

Stimuli were short auditory excerpts of 1.5 ± .2 sec, belonging to one of the following categories: Examples of the stimuli are available at http://personal.inb.unam.mx/lconcha/data.

2.3.1. *Human vocalizations*

This category included 1) ***Speech***, which consisted of 20 sentences spoken in several languages (e.g. Spanish, English, Mandarin; we include phrases in several languages and different speakers in order to eliminate possible bias attributable to the semantic content of the sentence or the gender of the speaker); and 2) ***Non-linguistic vocalizations***, consisting of 20 vocalizations such as yawning, laughs, and screams [A subset of these stimuli was previously used to identify voice-specific brain areas (Fecteau, Armony, Joanette, & Belin, 2004)]. Both subsets included female and male voices.

2.3.2. *Music*

Stimuli within this category were taken from longer musical passages that were unfamiliar and followed the rules of Western tonal music, expressing different emotions (Aubé, Angulo-Perkins, Peretz, Concha, & Armony, 2014; Vieillard et al., 2008). The musical excerpts (60 in total) were recorded with different instruments: 1) ***Piano***, (20 excerpts played by a pianist); 2) ***Violin***, (20 excerpts played by a violinist) and 3) ***Synthetic piano***, (20 excerpts produced by a computer with the timbre of a piano).

2.3.3. *Non-vocal sounds*

This category included 40 typical sounds produced by humans and representative of our everyday environment, such as starting a car, closing doors, honking, toilet flushing, etc. These stimuli were acoustically rich and easily recognizable by the volunteers.

2.3.4. *Monkey vocalizations*

20 monkey vocalizations were included as a control condition (Fecteau et al., 2004) as these stimuli have similar acoustic properties as human vocalizations but have no semantic or affective meaning to most humans (Fecteau, Belin, Joanette, & Armony, 2007).

2.3.5. *Silence*

Blocks of silence (i.e., scanner noise alone) were included as baseline.

Acoustical features were computed using MIRToolbox (Lartillot & Toiviainen, 2007) implemented in MATLAB (Mathworks, Natick, MA) and are shown in Supplementary Table 1. In general, the majority of acoustical features were similar between categories, with the exception of piano, particularly for articulation, spectral centroid, and spectral flux. Spectro-temporal analysis of the stimuli showed that music (from all three sub-categories), speech, and human vocalizations showed similar patterns of modulation, while non-vocal sounds and monkey vocalizations showed no particular modulation peak in the spectro-temporal modulation plots (Supplementary Fig. 1).

2.4. *Image acquisition*

All images were acquired using a 3T MR750 scanner (General Electric, Waukesha, Wisconsin). Functional volumes consisted of 50 slices (3 mm thick), acquired with a gradient-echo, echo planar imaging sequence with the following parameters: FOV = 256 × 256 mm^2, matrix = 128 × 128 (thus yielding a voxel size = 2 × 2 × 3 mm^3), TR = 3000 msec, TE = 40 msec. Images were acquired with a 32-channel coil using parallel imaging with an acceleration factor of 2. A 3D T1-weighted volume was also acquired for registration purposes (resolution of 1 × 1 × 1 mm^3, TR = 2.3 sec, TE = 3 msec).

2.5. *Image processing and statistical analyses*

All image processing was carried out using *fsl* tools (fMRIB, Oxford UK); functional image analysis was conducted using FEAT (FMRI Expert Analysis Tool) version 5.98 and statistical analyses based on the general linear model. Higher-level analyses were performed using FLAME (FMRIB's Local Analysis of Mixed Effects) (Beckmann, Jenkinson, & Smith, 2003; Woolrich, Behrens, Beckmann, Jenkinson, & Smith, 2004). Single-subject first level analysis were performed using a design matrix which consisted of 7 explanatory variables (EV's) representing each of the categories presented in the experiment (i.e., violin, piano, synthetic piano, speech, non-linguistic vocalizations, non-vocal sounds and monkey vocalizations). For each subject we performed a fixed-effects model in order to analyze acoustic categories that encompass two or more EVs (e.g., the category "music" includes piano, synthetic piano and violin). Subject-specific contrasts were computed for each of the comparisons of interest (see below) and entered in random effects, between-subjects model (including musical expertise as a group factor). Images were registered to standard space MNI152. Correction for multiple comparisons was carried out using random field theory (voxel $z > 2.3$, cluster $p < .05$) unless otherwise specified.

2.6. *Analyses*

2.6.1. *Analysis 1*

In order to identify cortical regions that are involved in the processing of both music and human vocalizations, we conducted a conjunction analysis to compare these sounds against two other categories of sounds that are also acoustically complex: non-vocal sounds generated by man and his daily human environment (e.g. footsteps, honking of a car) and non-human vocal sounds, such as monkey vocalizations. To find the common regions that responded to both music (including piano, violin and synthetic piano) and human vocalizations (including speech and non-linguistic vocalizations), compared to non-vocal sounds, we conducted a conjunction analysis (Nichols, Brett, Andersson, Wager, & Poline, 2005). This analysis identified voxels that were significantly activated (according to the statistical threshold mentioned above) in both music versus non-vocal sounds and human vocalizations versus non-vocal sounds contrasts. A second, similar conjunction was conducted for the comparison of music and human voice versus monkey vocalizations.

2.6.2. *Analysis 2*

The purpose of this analysis was to highlight "music-selective" regions, i.e., those regions activated predominantly by musical stimuli (piano, violin, and synthetic piano) in comparison to human vocal sounds (speech and non-linguistic vocalizations). The following contrasts allowed us to subtract functional regions that are shared between stimuli (probably the majority), and leave only those regions that are activated significantly more during music perception than by human vocal sounds. The analysis consisted of three main comparisons: [a] Music (piano, violin and synthetic piano) versus Human vocalizations (speech and non-linguistic vocalizations); [b] Music (piano, violin and synthetic piano) versus Speech; and [c] Music (piano, violin and synthetic piano) versus Non-linguistic vocalizations. To further assess the specificity of the putative music areas and to minimize potential confounding effects due to differences in basic acoustic properties between categories, we conducted a subsidiary analysis comparing violin versus speech, as these categories are acoustically more similar (Supplementary Fig. 1). In these analyzes we included all 53 subjects, with musicianship modeled as a covariate (Beckmann et al., 2003).

To evaluate the robustness of the main contrast testing music versus human vocalizations at the individual level, we visually examined each subject's corresponding statistical map and recorded if that particular subject showed activations within the *planum polare* reminiscent of the group result, using a threshold established at $p < .01$ (uncorrected). We performed a chi-square test to evaluate whether the observed activations were different between the two groups. The corresponding subject-level statistical maps (in standard space) thresholded at $p < .01$ (uncorrected) were combined to generate voxel-wise activation prevalence maps.

To further analyze the patterns of activation of the *planum polare* in response to human vocalizations and music, we extracted the percentage of BOLD signal change per subject using an atlas-based anatomical region of interest (ROI), drawn from the Harvard–Oxford Probabilistic Anatomical Atlas for each hemisphere (thresholded at 33%; resulting in volumes of 2,240 and 2,304 mm^3 that overlapped with the clusters resulting from Analysis 2 in 1,600 and 1,080 mm^3 for right and left hemispheres, respectively). The use of an atlas-based ROI ensures an unbiased examination of the fMRI data, as it avoids the pre-selection of voxels to be analyzed based on a prior result (Kriegeskorte, Simmons, Bellgowan, & Baker, 2009). We performed a mixed-design ANOVA (analysis of variance) on these data with hemisphere and stimulus category as within-subject factors and group (musicians versus non-musicians) as a between-subjects factor. Effects were considered significant at $p < .05$ significance level.

2.6.3. *Analysis 3*

In this analysis we compared musicians versus non-musicians in order to identify activations related to musical expertise. To achieve this, we performed a two-sample *t*-test for the contrast music (piano, violin, synthetic piano) versus human vocalizations (speech + non-linguistic) in musicians greater than non-musicians. The opposite contrast (i.e. human vocalizations > music, in musicians greater than non-musicians) was performed to evaluate whether auditory expertise in music processing generates functional changes in the perception of human vocalizations. As in Analysis 2, we analyze the variations of BOLD signal change using an anatomical ROI corresponding to the *planum temporale* (Harvard–Oxford probabilistic anatomical atlas, thresholded at 33%; which overlapped with the functional cluster in a region of 536 mm^3) to compare the activity elicited by music and human vocalizations. A mixed-design ANOVA was used to examine possible interactions between factors (group, acoustic category and hemisphere).

Finally, we correlated the behavioral information provided by the musicians related to the number of hours of training per week, starting age, years of experience and musical diversity (number of instruments played) with BOLD signal changes (relative to baseline) elicited by music listening.

3. Results

To ensure that each sound category produced significant activations in auditory regions in the presence of the scanner noise, we first contrasted each of them with baseline (silence blocks). All stimulus categories produced strong cortical activations within and beyond the auditory cortices that exceeded the activation produced by scanner noise alone. These results showed a large overlap in almost all the temporal regions activated, not only between music and speech, but also among the other stimuli (Fig. 2, panel A). All subjects responded appropriately to the target (overall response rate was 94.7% and false alarms rate was .05%).

3.1. *Analysis 1: common regions between music and human vocalizations*

3.1.1. *Human vocalizations and music > non-vocal sounds*

This analysis revealed activity occupying the anterior portion of the superior temporal gyrus (aSTG) in both hemispheres (left cluster = 4,756 mm^3; right cluster = 4,720 mm^3). The cluster distribution covered a large portion of the lateral face of the aSTG, including the *planum polare* and extended toward the temporal pole (Fig. 2, panel B). No region showed the opposite relationship (**Non-Vocal sounds > Human Vocalizations and Music**). Cluster volume and coordinates of peak significance for this and all subsequent contrasts are shown in Table 1.

3.1.2. *Human vocalizations and music > monkey vocalizations*

We found bilateral activation in the aSTG (right cluster = 5,328 mm^3; left cluster = 3,848 mm^3). Both clusters had distributions similar to those in the previous analysis but were slightly narrower (Fig. 2, panel C). The opposite contrast, testing for **Monkey vocalizations > Human Vocalizations and Music,** did not show significant activations.

3.2. *Analysis 2: cortical regions preferential to musical stimuli*

3.2.1. *Music (piano, violin and synthetic piano) > human vocalizations (speech and non-linguistic vocalizations)*

This contrast revealed a discrete and well-circumscribed bilateral area, located in the aSTG that responded significantly more to music than to human vocalizations (Fig. 3, cold colors). This region corresponds to the most anterior part of Brodmann's area 22 (BA 22) and to the posterior portion of BA 38, known as the *planum polare* (right cluster volume = 5,296 mm^3; left cluster volume = 3,144 mm^3). Outside the auditory cortices, we identified a cluster located on the upper portion of the left postcentral gyrus that corresponds to the primary somatosensory cortex.

The opposite contrast, **Human vocalizations > Music,** elicited bilateral activation of the lateral STG, medial temporal gyrus (MTG), and the superior temporal sulcus (STS). Specifically, in the left hemisphere we found a large cluster of activation occupying the lateral aspect of the STG and STS, and the cluster extended along the temporo-occipital region until the anterior division of the STG (Fig. 3, warm colors). The right hemisphere cluster had its peak of statistical significance located at coordinates very similar to those in the left hemisphere, but slightly more anterior and ventral. While the right hemisphere cluster was considerably smaller than that of the left hemisphere (23,408 *vs* 35,336 mm^3), their patterns of distribution were similar. Bilateral activation of the amygdala was found, and two clusters were evident in the frontal lobe: one located in the left inferior frontal gyrus (IFG), with its maximal peak of activation approximately in Broca's area (Brodmann areas 44 and 45), but with several local maxima with coordinates corresponding to the frontal orbital cortex; the second cluster was located in the medial portion of the left superior frontal gyrus (SFG).

3.2.2. *Music (piano, violin and synthetic piano) > speech*

No region reached the statistical threshold in the contrast testing for music (piano, violin, and synthetic piano) > speech after random field theory correction. However, using an

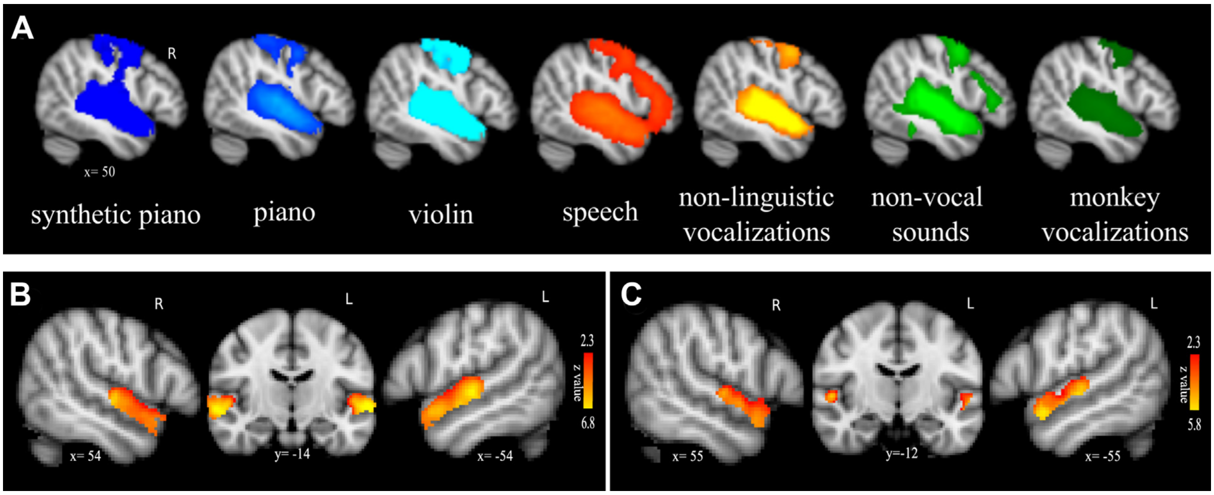

Fig. 2 – Global activity (above baseline) for all stimuli (panel A). Shared regions between music and human vocalizations in comparison to non-vocal sounds (panel B) and monkey vocalizations (panel C). Activations are overlaid on the MNI-152 atlas.

Table 1 – Significant activations for all experiments.

Contrast	Anatomical location	Peak			Cluster volume (mm^3)
		x	y	z	
Human vocalizations + music > non-vocal sounds	aSTG/temporal pole (left)	−64	−14	−4	9512
Human vocalizations + music > non-vocal sounds	aSTG/temporal pole (right)	62	−15	−4	9440
Human vocalizations + Music > monkey vocalizations	aSTG/temporal pole (right)	54	−12	2	5328
Human vocalizations + music > monkey vocalizations	aSTG/temporal pole (left)	−54	4	−14	3848
Music > human vocalizations	aSTG/*planum polare* (right)	50	−2	−8	5296
Music > Human vocalizations	aSTG/*planum polare* (left)	−48	−4	−8	3144
Music > human vocalizations	Primary somatosensory cortex (left)	−32	−32	48	7128
Human vocalizations > music	MTG/STG, STS (left)	−58	−8	12	35,336
Human vocalizations > music	MTG/STG, STS (right)	58	−6	−14	23,408
Human vocalizations > music	IFG/orbital cortex (left)	−52	20	−2	17,024
Human vocalizations > music	SFG middle portion, (left)	−8	62	24	3032
Human vocalizations > music	Temporal lobe, Amygdala (left)	−20	−6	−20	5192
Human vocalizations > music	Temporal lobe, Amygdala (right)	22	−2	−18	4616
Speech > music	MTG/STG, STS (left)	−64	−14	−4	59,624
Speech > music	MTG/pSTG, STS (right)	56	−18	8	42,216
Speech > music	IFG, (left)	−52	22	−8	19,856
Speech > music	IFG, (right)	52	28	14	2736
Speech > music	SFG middle portion, (left)	−4	40	42	5000
Speech > music	Primary motor cortex (left)	−50	−4	46	4048
Speech > music	Temporal lobe, Amygdala (left)	20	−8	−20	6504
Speech > music	Temporal lobe, Amygdala (right)	22	−2	−18	5264
Music > non-linguistic	STG (left)	−50	−6	−4	14,088
Music > non-linguistic	STG (right)	50	−4	−4	13,800
Music > non-linguistic	Primary somatosensory cortex (left)	−32	−32	48	6248
Violin > speech	aSTG/(right) *planum polare*	42	−12	−10	2696
Speech > violin	MTG/STG, STS (left)	−58	−18	−8	41,968
Speech > violin	MTG/pSTG, STS (right)	58	−8	−10	27,720
Speech > violin	IFG, (left)	−52	22	−4	18,784
Speech > violin	IFG, (right)	52	30	8	2712
Speech > violin	SFG middle portion, (left)	−4	40	42	4304
Speech > violin	Primary motor cortex (left)	−50	0	44	3328
Speech > violin	Temporal lobe, Amygdala (left)	−20	−10	−20	8824
Speech > violin	Temporal lobe, Amygdala (right)	22	−6	−18	7064
Music > human|musicians > non-musicians	*planum temporale* (right)	58	−14	4	4216

Human vocalizations = speech and non-linguistic vocalizations; Music = violin, piano, and synthetic piano; pMTG, posterior medial temporal gyrus; pSTG, posterior superior temporal gyrus; aMTG, anterior medial temporal gyrus; aSTG, anterior superior temporal gyrus; STS, superior temporal sulcus; IFG, inferior frontal gyrus; SFG, superior frontal gyrus. Cluster peaks are given in MNI coordinates in mm.

uncorrected *p*-value of .01, bilateral activation of the *planum polare* (BA 22/38) was evident. This bilateral activation is not different from that seen with the previous contrast (Fig. 3; dark blue color). With a *p*-value threshold of .001 (uncorrected) we still found 2% of the cluster located on the *planum polare*, bilaterally.

Functional maps testing for **Speech** > **Music,** showed regions similar to those found in the contrast testing for human vocalizations > music (Fig. 3; warm colors), bilateral activation of the lateral aspect of the STG extending to the STS and MTG, and bilateral activation of the IFG. However, it covered Brodmann areas 44 and 45 only in the left hemisphere, whereas in the right hemisphere it included only the most anterior part of Broca's area (BA 45). Other active regions included the amygdalae, the medial region of the SFG, and the left premotor cortex (data not shown).

3.2.3. *Music (piano, violin and synthetic piano) > non-linguistic vocalizations*

This comparison elicited bilateral activations, occupying a large part of the STG from the end of the posterior ramus of the STG (near Brodmann's area 37) to the most anterior portion of the gyrus (covering almost the full extent of Brodmann's area 22). The cluster distribution was slightly narrower in the right hemisphere than in the left hemisphere (Supplementary Fig. 2, left). As observed in the first contrast (music > human vocalizations), the dorsal portion of the left primary somatosensory cortex was activated.

The opposite contrast (**Non-linguistic vocalizations** > **Music**) did not yield statistically significant results.

3.2.4. *Violin > speech*

We found a cluster located in the *planum polare* of the right hemisphere that overlaps with the cluster found in all previous contrasts testing for music > human vocalizations; yet unlike those contrasts, the BOLD signal in its left-hemisphere counterpart did not reach statistical significance (Fig. 4; blue). The contrast analysis for **Speech** > **Violin** activated the same eight clusters described previously in the comparison of speech > music (Fig. 4, warm colors).

We also evaluated the activation generated by music in comparison to the other complex stimuli (non-vocal sounds

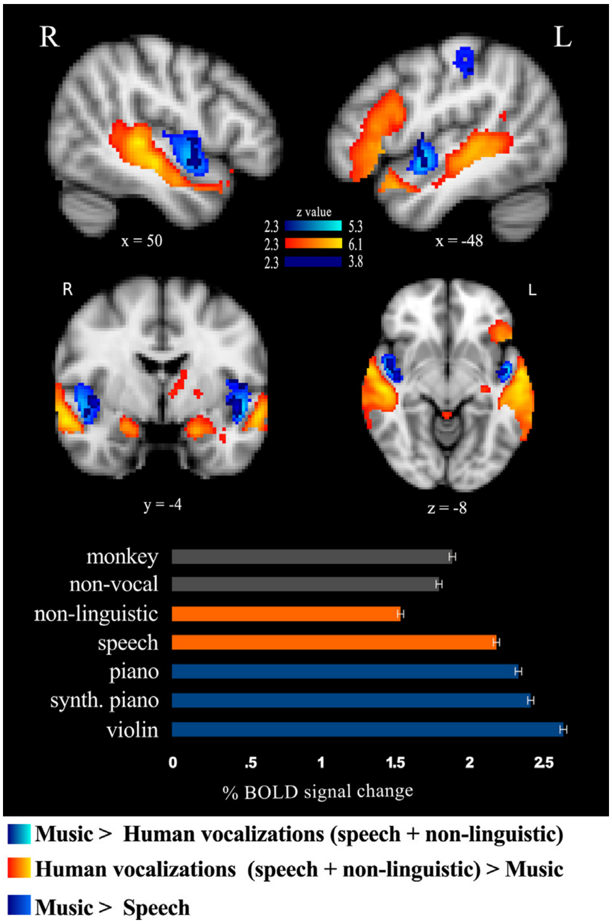

Fig. 3 – Differential patterns of activation for music and human vocalization stimuli. The clusters in light blue (music > human vocalizations [speech + non-linguistic]) and dark blue (music > speech) show no overlap with the cluster in orange (human vocalizations > music). Bottom panel: Group average (±SEM) of BOLD signal changes above baseline as a function of acoustic stimulus categories, taken from the peak of statistical significance of the cluster identified in the main analysis (music > human vocalizations, right *planum polare*, Table 1). All 53 subjects, regardless of musicianship, are included in these analyses.

and monkey vocalizations). The results are shown in the Supplementary Materials.

3.2.5. *Subject-level analyses*

Prevalence activation maps of the region found at the group level (*planum polare*) in the main contrast (music versus human vocalizations), showed that bilateral activation of the *planum polare* was highly prevalent in the musicians group (22 of 28 subjects), unlike what was observed in the non-musicians, where less than half had bilateral activation (8 of 25 subjects; χ^2, $p = .00064$) (Fig. 5, panel A). Likewise, 27 of the 28 musicians showed activity in the right hemisphere, while 13 of the 25 non-musicians presented it in the same hemisphere (χ^2, $p = .000175$). In the left hemisphere, focal activity within the *planum polare* was present in 23 of 28 musicians and in 13 of 25 non-musicians (χ^2, $p = .0189$). In the majority of the subjects analyzed, cortical activity within the *planum polare* was considerably greater in response to music listening than to human sounds. The mixed-design ANOVA of the BOLD signal change within the anatomical ROIs (left and right *planum polare*) showed a significant main effect for stimuli category ($p < .000003$) and hemisphere ($p < .000048$). There was a significant interaction between stimulus category and hemisphere ($p < .02$) but no interactions between group and stimulus category or hemisphere (Fig. 5, panel C).

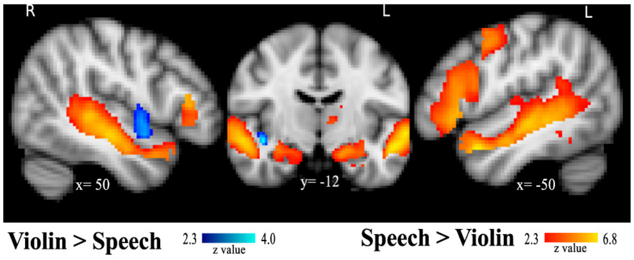

Fig. 4 – Violin versus speech. The *planum polare* shows significant activation in response to violin stimuli versus speech (light blue). The opposite contrast (i.e., speech > violin, warm colors) shows a more distributed network involving the temporal and frontal lobes.

3.3. *Analysis 3: musicians versus non-musicians*

3.3.1. *Music (piano, violin, synthetic piano) > human vocalizations (speech and non-linguistic vocalizations)*

We found an interaction between cortical activity and musicianship evidenced by a greater differential BOLD signal in the right *planum temporale* of the musicians. The cluster covered the lateral face of the posterior portion of the STG and was distributed medially until the limit of Heschl's gyrus. This region does not overlap with the *planum polare*, but is located postero-laterally to it (Fig. 6, green cluster). Examination of the BOLD signal change within the resulting cluster showed that speech elicited the largest signal changes regardless of musicianship (Fig. 6, lower right panel). However, we found significant differences between conditions (music and human vocalizations) only in the group of non-musicians, with greater activations elicited by human vocalizations than music ($p = 5.1 \times 10^{-7}$; Fig. 6, lower left panel). No significant differences were found between acoustic categories in musicians ($p = .95$).

No region surpassed the statistical threshold for the contrast testing for **Human vocalizations** > **Music.**

The mixed-design ANOVA conducted in the anatomical ROI corresponding to the *planum temporale*, revealed a main effect for stimulus category ($p < .006$) and an interaction effect between group and stimulus category ($p < .01$), with non-musicians showing higher activity in response to human vocalizations than musical stimuli.

3.3.2. *Correlation with musical expertise*

Based on the results outlined above, we investigated whether the BOLD activity of the *planum polare* correlated with musical experience. For each subject in the musician group we obtained the percent BOLD signal change produced by music

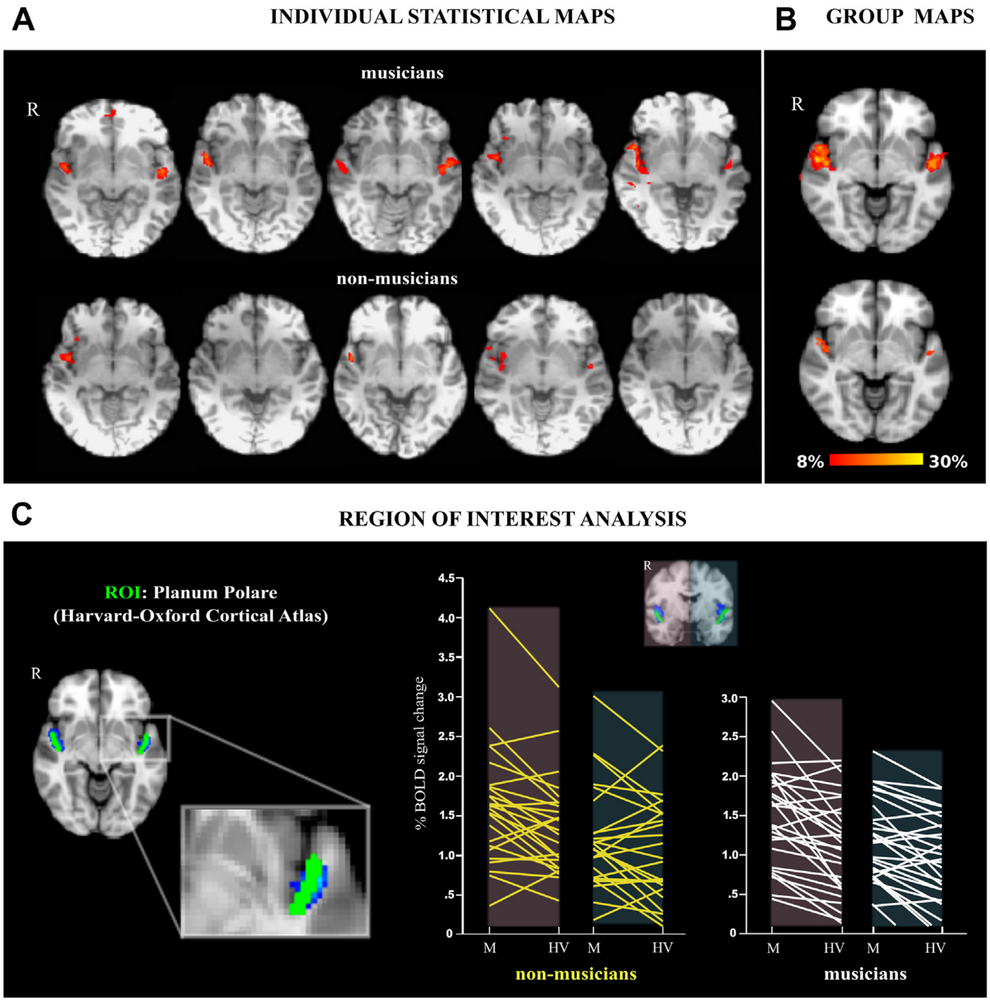

Fig. 5 – Subject-specific activations. A, individual statistical maps from the analyses for music > human vocalizations (red), showing T1-weighted images of five representative musicians (upper panel) and five representative non-musicians (lower panel) overlaid with maps of statistically significant parameter estimates of differential BOLD response (thresholded at $p < .01$, uncorrected). While activation of the right *planum polare* is common in both groups, bilateral activation is more prevalent in musicians. B, activation prevalence maps showing more overlap of the clusters in musicians than in non-musicians. C, region of interest analysis (ROI). The anatomical ROI corresponding to the *planum polare* (green, derived from the Harvard–Oxford cortical atlas) closely overlaps with our results (blue). Within this ROI, subject-specific percent BOLD signal changes in response to stimulation with music (M) and human vocalizations (HV) are shown.

listening within the voxel with the peak significance identified in Analysis 1 (Section 3.1) for the right and left *planum polare* clusters separately (cold colors in Fig. 3). We found that musicians' hours of practice per week showed a positive correlation with BOLD activation in the left ($r = .408$; $p = .03$), but not the right cluster. We did not find any significant correlations between BOLD activity of the *planum polare* and age at start of musical training, number of instruments played, or years of experience.

4. Discussion

In this paper, we describe the cortical regions involved in the processing of music and compare them with the brain areas associated with the perception of human vocal sounds (with or without linguistic content). Previous lesion studies have shown that speech and music, in terms of both comprehension and production, can be affected in a selective manner (e.g., *aphasia without amusia* or *amusia without aphasia*) (Peretz et al., 2002). However, other groups have found evidence suggesting that music and speech processing is shared (Abrams et al., 2011; Koelsch, 2005a; Patel, 2003). Our results are consistent, to some extent, with both views. We observed considerable overlap between the temporal regions activated by music and speech that is not shared by other complex stimuli, such as non-vocal sounds and monkey vocalizations. However, our analyses also revealed cortical regions that are preferentially activated by specific sound categories. Music (explored here by using violin, piano and synthetic piano melodies) in particular elicited significantly stronger activation than human vocalizations in the *planum polare*, an area

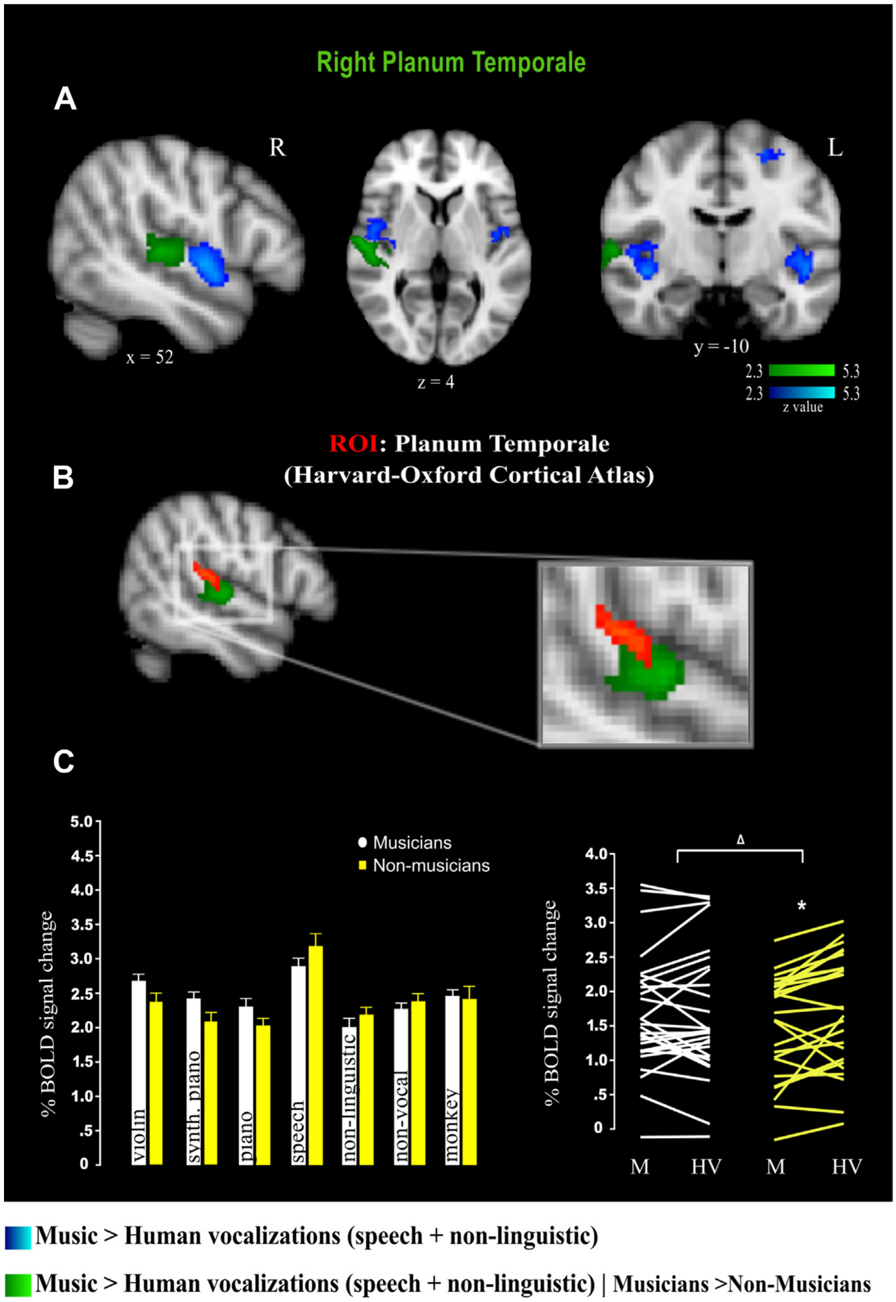

Fig. 6 – Differences between musicians and non-musicians. A: Differential BOLD activity elicited by music or human vocalizations was modulated by musicianship only within the right *planum temporale* (green cluster). Regions showing greater activity to music than human vocalizations regardless of musicianship (Fig. 2) are shown in blue, for reference. B: Anatomically-defined region of interest (ROI) corresponding to the *planum temporale* (red) taken from the Harvard–Oxford Cortical Atlas and overlapped with results from Analysis 3 (green). C (left): BOLD signal in the *planum temporale* during all stimuli (mean ± SEM, taken from the voxel with largest statistical significance for illustration purposes). For each group, speech produces higher activations than those elicited by any of the other acoustic stimuli, yet non-musicians show a much larger difference of activity between speech and musical stimuli than the group of musicians. C (right): Percentage of BOLD signal change in the anatomical ROI of the *planum temporale* elicited by music (M) and human vocalizations (HV) in musicians (white) and in non-musicians (yellow). Only non-musicians show statistically significant differences between acoustic categories (asterisk, $p < .005$). Delta denotes a significant interaction between group and acoustic category.

that is known to co-participate with frontal regions in high-level music tasks involving melody repetition and harmonization (Brown et al., 2006) as well as pitch and/or melodic discrimination (Koelsch et al., 2002). Likewise, the aSTG has been reported to be a belt/parabelt area sensitive to complex spectral changes, in contrast to primary or core regions which respond prominently to pure tones (Woods et al., 2009, 2010; Zatorre & Belin, 2001).

Our results show that fMRI offers sufficient spatial resolution to reveal category-specific specialization of the temporal cortex, and they are consistent with and extend recent studies that report differences in the networks involved in analyzing sentences or musical stimuli (Fedorenko, McDermott, Norman-Haignere, & Kanwisher, 2012; Lai et al., 2012; Rogalsky et al., 2011). For instance, Rogalsky et al. (2011) used multivariate pattern classification analysis and a sound stimulation paradigm with music and speech stimuli, both artificially modulated in their envelope rate, in order to isolate higher-level processing, and they reported no overlap between the regions activated (activation of the dorsomedial STG for music, and the ventrolateral STG when the stimulus was a sentence). In addition, Lai et al. (2012) studied a population of children with autism, and also reported greater activation in the *planum polare* during music listening than in speech listening (songs and sentences sung or spoken by their parents) in both their control group and in autistic children; the opposite contrast (i.e., speech > music) produced an activation similar to what we found but only in their control group, with autistic children not showing this differential activation. Likewise, Tierney, Dick, Deutsch, and Sereno (2013) reported an increased response in the aSTG (*planum polare*) during a musical illusion paradigm in which the illusory perception of a melody is suddenly formed out of acoustic elements contained in a spoken phrase that was repeated constantly. Their finding suggests that the *planum polare* is involved not only in the spectro-temporal decomposition of sound, but that it may play an active and important role in the perception of sound categories, particularly in music.

The relatively large number of subjects included in our study allowed us to examine, in each group separately, the activation of the cortical networks during music listening at the individual level. We observed that musical stimuli elicited stronger activations of the *planum polare* than those in response to human vocalizations. Interestingly, this difference was seen in the left *planum polare* regardless of musicianship, but bilateral activation of this region was more prevalent in the musician group (Fig. 5; panel B). While we cannot attribute musical expertise as a defining factor in how music and human vocalizations are processed within the temporal lobes, as we did not find a significant interaction between group and hemisphere (or stimulus category), our data revealed a positive correlation between the activation level of the *planum polare* during music listening and hours of practice per week.

We found a single cluster (*planum temporale*), in which musicianship modulated the differential activity elicited by music and human vocalizations (Fig. 6). The BOLD signal change analysis performed in this region using an anatomical ROI revealed a main effect of stimulus category (proving its sensitivity to speech [Belin, Zatorre, Lafaille, Ahad, & Pike, 2000, Belin, Zatorre, & Ahad, 2002]), but also an interaction effect between group and stimulus type: whereas non-musicians showed considerably higher activity of the *planum temporale* in response to speech (as compared to music), musicians showed similar responses for both types of acoustic categories. If we consider that musicianship plays a role in the recruitment of specific areas within the temporal lobes, it is possible that, in musicians, these regions contribute to the analysis of musical components in a more elaborate fashion, in relation to the auditory environment to which they are exposed. Nonetheless, our results cannot establish whether this extended network for music processing is a cause or a consequence of the musicians' skills.

The somatosensory cortex was also activated during music listening (compared with human vocalizations, speech, and non-vocal sounds), demonstrating cross-modal integration influenced by music listening. While the role of the somatosensory cortices during music listening is unclear, other groups have reported that musical training changes the anatomical and functional properties of motor and somatosensory cortices (D'Ausilio, Altenmüller, Olivetti Belardinelli, & Lotze, 2006; Pantev & Herholz, 2011; Popescu, Otsuka, & Ioannides, 2004). Those studies suggest that musicians perform motor planning in response to music, perhaps in an attempt to produce or extend what they are listening to. We, however, did not find differences of activation attributable to musicianship.

Processing of human vocalizations recruit specific frontal regions, several of which have been reported since the first studies related to speech (Broca, 1863, 1865). Indeed, we found activity generated by human vocalizations outside the auditory cortices, localized in the hippocampus, amygdala, and IFG, all of which have been considered an essential part of the speech-processing network (Passingham, 1981), and have been studied in more appropriate stimulation paradigms for mnemonic (formation, recovery, and endurance of the memory trace) and emotional (perception, regulation, and inhibition) analyses of auditory processing (Fecteau, Belin, Joanette, & Armony, 2007; Henke, Buck, Weber, & Wieser, 1997; Koelsch, 2005b; Phelps & LeDoux, 2005; Watanabe, Yagishita, & Kikyo, 2008).

Our data agree with other neuroimaging studies pointing out that regions distal to the primary auditory cortex (e.g., *planum temporale* or *polare*) are involved in higher order acoustic analysis, unlike the activity of the primary auditory cortex (core), which is modulated by basic acoustic attributes such as frequency, intensity, or location (Woods et al., 2010). The *planum polare*, pSTG, and the lateral aspect of right Heschl's gyrus are activated by listening to sequences of variable pitch (e.g., tonal melody) (Patterson, Uppenkamp, Johnsrude, & Griffiths, 2002). Our choice of acoustical stimuli enabled us to compare different types of sounds without artificially matching their audio properties in order to retain their ecological validity and familiarity (Staeren, Renvall, De Martino, Goebel, & Formisano, 2009). Acoustical features, however, may play an important role in the differential activation of specific cortical areas beyond the primary auditory cortex. We investigated this in a separate analysis using data from a different fMRI paradigm that was acquired during the same session in all subjects included in this study. The paradigm was designed as an event-related experiment, tailored to address emotional features of sound, and it included stimuli similar to the ones outlined in the Methods (Section 2), but only for the music, speech, and non-linguistic categories. Acoustical parameters were extracted from these stimuli and correlated with the degree of cortical activation they induced within the *planum polare* bilaterally in the event-related experiment data (Aubé, Angulo-Perkins, Peretz, Concha, & Armony, 2014). We found significant correlations between

BOLD signal and articulation ($r = -.33$, $p < .0001$) as well as rhythmic clarity ($r = .35$, $p < .0001$). Both of these features are related to temporal aspects of sound, namely the way two sounds are connected (i.e., the form of the transition between them) and how perceptible a beat is within a musical fragment (Friberg & Hedblad, 2011), suggesting that the *planum polare* may be involved in the processing of specific temporal characteristics that are inherent to music.

Taken together, our data suggest that even though there are regions within the temporal lobes which are more responsive to both music and human vocalizations, compared to human non-vocal sounds and non-human primate vocalizations, specific functional regions exist, showing greater sensitivity to a particular sound category, with music-specific cortical networks likely involved in higher order analyses, such as extraction of melodic information (Patterson et al., 2002; Zatorre, Belin, & Penhune, 2002). Further, musicianship and extensive training seem to modulate the activity elicited by music and voice in the right *planum temporale*, suggesting functional plasticity in response to environmental exposure. These results contribute to a better functional characterization of the auditory cortex and, in particular, they describe the central role of the *planum polare* for music perception, providing a basis for future opportunities to design hearing, music, and speech therapies.

Acknowledgments

This study was sponsored by Conacyt (IE252-120295) and UNAM-DGAPA (I1202811 and IN212811). We thank the Posgrado en Ciencias Biomédicas of the Universidad Nacional Autónoma de México (UNAM) and the CONACyT for Graduate Fellowship 233109 to A.P.A. We are grateful to Marc Schönwiesner for his help with the spectro-temporal analysis of our acoustic stimuli; to Marc Hauser for providing the monkey vocalization recordings; to Pascal Belin and Shirley Fecteau for providing some of the non-linguistic vocalizations; to José Luis Díaz and Hugo Merchant for their insightful comments; to Daniel Ramírez Peña, Juan Ortíz-Retana, Erick Pasaye, and Leopoldo González-Santos for technical assistance during data acquisition and analysis; and to Dorothy Pless for proof-reading and editing our manuscript. We also thank the authorities of the Music Conservatory "José Guadalupe Velázquez", the Querétaro Philharmonic Orchestra, the Querétaro School of Violin Making, and the Querétaro Fine Arts Institute, for their help with the recruitment of musicians. The authors declare no competing financial interests.

Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.cortex.2014.07.013.

REFERENCES

Abrams, D. A., Bhatara, A., Ryali, S., Balaban, E., Levitin, D. J., & Menon, V. (2011). Decoding temporal structure in music and speech relies on shared brain resources but elicits different fine-scale spatial patterns. *Cerebral Cortex, 21*, 1507–1518.

Aubé, W., Angulo-Perkins, A., Peretz, I., Concha, L., & Armony, J. L. (2014). Fear across the senses: brain responses to music, vocalizations and facial expressions. *Social Cognitive and Affective Neuroscience nsu067*. http://dx.doi.org/10.1093/scan/nsu067.

Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2003). General multilevel linear modeling for group analysis in fMRI. *NeuroImage, 20*, 1052–1063.

Belin, P., Zatorre, R. J., & Ahad, P. (2002). Human temporal-lobe response to vocal sounds. *Brain Research. Cognitive Brain Research, 13*, 17–26.

Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature, 403*, 309–312.

Broca, P. (1863). Localization des fonctions cérébrales: Siege du langage articulé. *Bulletin de la Societé d'Anthropologie, 4*, 200–203.

Broca, P. (1865). Du siegé de la faculté du langage articulé. *Bulletin de la Societé d'Anthropologie, 6*, 337–339.

Brown, S., Martinez, M. J., & Parsons, L. M. (2006). Music and language side by side in the brain: a PET study of the generation of melodies and sentences. *European Journal of Neuroscience, 23*, 2791–2803.

D'Ausilio, A., Altenmüller, E., Olivetti Belardinelli, M., & Lotze, M. (2006). Cross-modal plasticity of the motor cortex while listening to a rehearsed musical piece. *The European Journal of Neuroscience, 24*, 955–958.

Dick, F., Lee, H. L., Nusbaum, H., & Price, C. J. (2011). Auditory-motor expertise alters "speech selectivity" in professional musicians and actors. *Cerebral Cortex, 21*, 938–948.

Elmer, S., Hänggi, J., Meyer, M., & Jäncke, L. (2014). Increased cortical surface area of the left *planum temporale* in musicians facilitates the categorization of phonetic and temporal speech sounds. *Cortex, 49*(10), 2812–2821. http://dx.doi.org/10.1016/j.cortex.2013.03.007.

Fecteau, S., Armony, J. L., Joanette, Y., & Belin, P. (2004). Is voice processing species-specific in human auditory cortex? an fMRI study. *NeuroImage, 23*, 840–848.

Fecteau, S., Belin, P., Joanette, Y., & Armony, J. L. (2007). Amygdala responses to nonlinguistic emotional vocalizations. *NeuroImage, 36*, 480–487.

Fedorenko, E., McDermott, J. H., Norman-Haignere, S., & Kanwisher, N. (2012). Sensitivit to musical structure in the human brain. *Journal of Neurophysiology, 108*, 3289–3300.

Friberg, A., & Hedblad, A. (2011). A comparison of perceptual ratings and computed audio features. In *Proc. 8th Sound and Music Computing Conference* (pp. 122–127).

Harley, E. M., Pope, W. B., Villablanca, J. P., Mumford, J., Suh, R., Mazziotta, J. C., et al. (2009). Engagement of fusiform cortex and disengagement of lateral occipital cortex in the acquisition of radiological expertise. *Cerebral Cortex, 19*, 2746–2754.

Henke, K., Buck, A., Weber, B., & Wieser, H. G. (1997). Human hippocampus establishes associations in memory. *Hippocampus, 7*, 249–256.

Herdener, M., Esposito, F., di Salle, F., Boller, C., Hilti, C. C., Habermeyer, B., et al. (2010). Musical training induces functional plasticity in human hippocampus. *Journal of Neuroscience, 30*, 1377–1384.

Koelsch, S. (2005a). Neural substrates of processing syntax and semantics in music. *Current Opinion in Neurobiology, 15*, 207–212.

Koelsch, S. (2005b). Investigating emotion with music: neuroscientific approaches. *Annals of the New York Academy of Sciences, 1060*, 412–418.

Koelsch, S., Gunter, T. C., V Cramon, D. Y., Zysset, S., Lohmann, G., & Friederici, A. D. (2002). Bach speaks: a cortical "language-network" serves the processing of music. *NeuroImage, 17*, 956–966.

Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S., & Baker, C. I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nature Neuroscience, 5*, 535–540.

Lai, G., Pantazatos, S. P., Schneider, H., & Hirsch, J. (2012). Neural systems for speech and song in autism. *Brain, 135*, 961–975.

Lartillot, O., & Toiviainen, P. (2007). MIR in Matlab (II): a toolbox for musical feature extraction from audio. In *Proceedings of 8th International Conference on Music Information Retrieval*.

Masataka, N. (2009). The origins of language and the evolution of music: a comparative perspective. *Physics of Life Reviews, 6*, 11–22.

Nichols, T., Brett, M., Andersson, J., Wager, T., & Poline, J. B. (2005). Valid conjunction inference with the minimum statistic. *NeuroImage, 25*, 653–660.

Pantev, C., & Herholz, S. C. (2011). Plasticity of the human auditory cortex related to musical training. *Neuroscience and Biobehavioral Reviews, 35*, 2140–2154.

Passingham, R. E. (1981). Broca's area and the origins of human vocal skill. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 292*, 167–175.

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience, 6*, 674–681.

Patel, A. D. (2007). *Music, language, and the brain*. New York: Oxford University Press.

Patterson, R. D., Uppenkamp, S., Johnsrude, I. S., & Griffiths, T. D. (2002). The processing of temporal pitch and melody information in auditory cortex. *Neuron, 36*, 767–776.

Peretz, I. (2002). Brain specialization for music. *The Neuroscientist, 8*, 372–380.

Peretz, I. (2006). The nature of music from a biological perspective. *Cognition, 100*, 1–32.

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience, 6*, 688–691.

Phelps, E. A., & LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron, 48*, 175–187.

Pinker, S. (1999). *How the mind works*. New York: W. W. Norton & Company.

Popescu, M., Otsuka, A., & Ioannides, A. A. (2004). Dynamics of brain activity in motor and frontal cortical areas during music listening: a magnetoencephalographic study. *NeuroImage, 21*, 1622–1638.

Rogalsky, C., Rong, F., Saberi, K., & Hickok, G. (2011). Functional anatomy of language and music perception: temporal and structural factors investigated using functional magnetic resonance imaging. *Journal of Neuroscience, 31*, 3843–3852.

Schlaug, G., Jäncke, L., Huang, Y., Staiger, J. F., & Steinmetz, H. (1995). Increased corpus callosum size in musicians. *Neuropsychologia, 33*, 1047–1055.

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

Schmithorst, V. J. (2005). Separate cortical networks involved in music perception: preliminary functional MRI evidence for modularity of music processing. *NeuroImage, 25*, 444–451.

Staeren, N., Renvall, H., De Martino, F., Goebel, R., & Formisano, E. (2009). Sound categories are represented as distributed patterns in the human auditory cortex. *Current Biology:CB, 19*, 498–502.

Tierney, A., Dick, F., Deutsch, D., & Sereno, M. (2013). Speech versus song: multiple pitch-sensitive areas revealed by a naturally occurring musical illusion. *Cerebral Cortex, 23*, 249–254.

Vieillard, S., Peretz, I., Gosselin, N., Khalfa, S., Gagnon, L., & Bouchard, B. (2008). Happy, sad, scary and peaceful musical excerpts for research on emotions. *Cognition & Emotion, 22*, 720–752.

Wan, C. Y., & Schlaug, G. (2010). Music making as a tool for promoting brain plasticity across the life span. *The Neuroscientist, 16*, 566–577.

Watanabe, T., Yagishita, S., & Kikyo, H. (2008). Memory of music: roles of right hippocampus and left inferior frontal gyrus. *NeuroImage, 39*, 483–491.

Woods, D. L., Herron, T. J., Cate, A. D., Yund, E. W., Stecker, G. C., & Rinne, T. (2010). Functional properties of human auditory cortical fields. *Frontiers in Systems Neuroscience, 4*, 155.

Woods, D. L., Stecker, G. C., Rinne, T., Herron, T. J., Cate, A. D., Yund, E. W., et al. (2009). Functional maps of human auditory cortex: effects of acoustic features and attention. *PLoS One, 4*, e5183.

Woolrich, M. W., Behrens, T. E. J., Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2004). Multilevel linear modelling for FMRI group analysis using Bayesian inference. *NeuroImage, 21*, 1732–1747.

Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory corte. *Cerebral Cortex, 11*, 946–953.

Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends in Cognitive Sciences, 6*, 37–46.