

Sensitivity to musical structure in the human brain

Evelina Fedorenko,¹ Josh H. McDermott,² Sam Norman-Haignere,¹ and Nancy Kanwisher¹

¹Department of Brain and Cognitive Sciences, MIT, Cambridge, Massachusetts; and ²Center for Neural Science, New York University, New York, New York

Submitted 9 March 2012; accepted in final form 23 September 2012

Fedorenko E, McDermott JH, Norman-Haignere S, Kanwisher N. Sensitivity to musical structure in the human brain. *J Neurophysiol* 108: 3289–3300, 2012. First published September 26, 2012; doi:10.1152/jn.00209.2012.—Evidence from brain-damaged patients suggests that regions in the temporal lobes, distinct from those engaged in lower-level auditory analysis, process the pitch and rhythmic structure in music. In contrast, neuroimaging studies targeting the representation of music structure have primarily implicated regions in the inferior frontal cortices. Combining individual-subject fMRI analyses with a scrambling method that manipulated musical structure, we provide evidence of brain regions sensitive to musical structure bilaterally in the temporal lobes, thus reconciling the neuroimaging and patient findings. We further show that these regions are sensitive to the scrambling of both pitch and rhythmic structure but are insensitive to high-level linguistic structure. Our results suggest the existence of brain regions with representations of musical structure that are distinct from high-level linguistic representations and lower-level acoustic representations. These regions provide targets for future research investigating possible neural specialization for music or its associated mental processes.

brain; fMRI; music

MUSIC IS UNIVERSALLY and uniquely human (see, e.g., McDermott and Hauser 2005; Stalinski and Schellenberg 2012; Stevens 2012). A central characteristic of music is that it is governed by structural principles that specify the relationships among notes that make up melodies and chords and beats that make up rhythms (see, e.g., Jackendoff and Lerdahl 2006; Krumhansl 2000; Tillmann et al. 2000 for overviews). What mechanisms in the human brain process these structural properties of music, and what can they tell us about the cognitive architecture of music?

Some of the earliest insights about high-level musical processing came from the study of patients with brain damage. Damage to temporal lobe structures (often in the right hemisphere; Milner 1962) can lead to “amusia,” a deficit in one or more aspects of musical processing (enjoying, recognizing, and memorizing melodies or keeping rhythm), despite normal levels of general intelligence and linguistic ability (see, e.g., Peretz and Coltheart 2003; Peretz and Hyde 2003). Critically, some patients with musical deficits demonstrate relatively preserved lower-level perceptual abilities, such as that of discriminating pairs or even short sequences of tones (e.g., Allen 1878; Di Pietro et al. 2004; Griffiths et al. 1997; Liegeois-Chauvel et al. 1998; Patel et al. 1998b; Peretz et al. 1994; Phillips-Silver et al. 2011; Piccirilli et al. 2000; Steinke et al. 2001; Stewart et al. 2006; Warrier and Zatorre 2004; Wilson et al. 2002). Perhaps the most striking case is that of patient G.L. (Peretz et al. 1994), who—following damage to

left temporal lobe and fronto-opercular regions—could judge the direction of note-to-note pitch changes and was sensitive to differences in melodic contour in short melodies, yet was unable to tell the difference between tonal and atonal musical pieces or make judgments about the appropriateness of a note in a musical context, tasks that are trivial for most individuals even without musical training (e.g., Bharucha 1984; Dowling and Harwood 1986). These findings suggest that mechanisms beyond those responsible for basic auditory analysis are important for processing structure in music.

Consistent with these patient studies, early brain imaging investigations that contrasted listening to music with low-level baselines like silence or noise bursts reported activations in the temporal cortices (e.g., Binder et al. 2000; Evers et al. 1999; Griffiths et al. 1999; Patterson et al. 2002; Zatorre et al. 1994). However, neuroimaging studies that later attempted to isolate structural processing in music (distinct from generic auditory processing) instead implicated regions in the frontal lobes. Two key approaches have been used to investigate the processing of musical structure with fMRI: 1) examining responses to individual violations of musical structure (e.g., Koelsch et al. 2002, 2005; Tervaniemi et al. 2006; Tillmann et al. 2006), using methods adopted from the event-related potential (ERP) literature (e.g., Besson and Faita 1995; Janata 1995; Patel et al. 1998a), and 2) comparing responses to intact and “scrambled” music (e.g., Abrams et al. 2011; Levitin and Menon 2003, 2005). Violation studies have implicated posterior parts of the inferior frontal gyrus (IFG), “Broca’s area” (e.g., Koelsch et al. 2002; Maess et al. 2001; Sammler et al. 2011), and scrambling studies have implicated the more anterior, orbital, parts of the IFG in and around Brodmann area (BA) 47 (e.g., Levitin and Menon 2003). Although the violations approach has high temporal precision and is thus well suited for investigating questions about the time course of processing musical structure, such violations sometimes recruit generic processes that are engaged by irregularities across many different domains. For example, Koelsch et al. (2005) demonstrated that all of the brain regions that respond to structural violations in music also respond to other auditory manipulations, such as unexpected timbre changes (see also Doeller et al. 2003; Opitz et al. 2002; Tillmann et al. 2003; see Corbetta and Shulman 2002 for a meta-analysis of studies investigating the processing of low-level infrequent events that implicates a similar set of brain structures; cf. Garza Villarreal et al. 2011; Koelsch et al. 2001; Leino et al. 2007). We therefore chose to use a scrambling manipulation in the present experiment.

Specifically, we searched for regions that responded more strongly to intact than scrambled music, using a scrambling procedure that manipulated musical structure by randomizing the

Address for reprint requests and other correspondence: E. Fedorenko, MIT, 43 Vassar St., 46-3037G, Cambridge, MA 02139 (e-mail: evelina9@mit.edu).

pitch and/or timing of each note.¹ We then asked 1) whether any of these regions are located in the temporal lobes (as implicated in prior neuropsychological studies), 2) whether these regions are sensitive to pitch scrambling, rhythm scrambling, or both, and 3) whether these regions are also responsive to high-level linguistic structure² (i.e., the presence of syntactic and semantic relationships among words). Concerning the latter question, a number of ERP, magnetoencephalography (MEG), fMRI, and behavioral studies have argued for overlap in processing musical and linguistic structure (e.g., Fedorenko et al. 2009; Hoch et al. 2011; Koelsch et al. 2002, 2005; Maess et al. 2001; Patel et al. 1998a; Slevc et al. 2009; see, e.g., Koelsch 2005; Slevc 2012; or Tillmann 2012 for reviews), but double-dissociations in patients suggest at least some degree of independence (e.g., Dalla Bella and Peretz 1999; Luria et al. 1965; Peretz 1993; Peretz and Coltheart 2003). Consistent with the patient studies, two recent fMRI studies found little response to music in language-structure-sensitive brain regions (Fedorenko et al. 2011; Rogalsky et al. 2011). However, to the best of our knowledge, no previous fMRI study has examined the response of music-structure-sensitive brain regions to high-level linguistic structure. Yet such regions are predicted to exist by the patient evidence (e.g., Peretz et al. 1994). We addressed these research questions by using analysis methods that take into account anatomical and functional variability (Fedorenko et al. 2010; Nieto-Castañón and Fedorenko 2012), which is quite pronounced in the temporal lobe (e.g., Frost and Goebel 2011; Geschwind and Levitsky 1968; Keller et al. 2007; Nieto-Castañón et al. 2003; Ono et al. 1990; Pernet et al. 2007; Tahmasebi et al. 2012).

METHODS

Participants. Twelve participants (6 women, 6 men) between the ages of 18 and 50 yr—students at MIT and members of the surrounding community—were paid for their participation. Participants were right-handed native speakers of English without extensive musical training (no participant had played a musical instrument for an extended period of time; if a participant took music lessons it was at least 5 yr prior to the study and for no longer than 1 yr). All participants had normal hearing and normal or corrected-to-normal vision and were naive to the purposes of the study. All protocols were reviewed and approved by the Internal Review Board at MIT, and all participants gave informed consent in accordance with the requirements of the Internal Review Board. Four additional participants were scanned but not included in the analyses because of excessive motion, self-reported sleepiness, or scanner artifacts.

Design, materials, and procedure. Each participant was run on a music task and then a language task. The entire scanning session lasted between 1.5 and 2 h.

Music task. There were four conditions: Intact Music, Scrambled Music, Pitch Scrambled Music, and Rhythm Scrambled Music. Each condition was derived from musical instrument digital interface (MIDI) versions of unfamiliar pop/rock music from the 1950s and 1960s. (The familiarity of the musical pieces was assessed informally

by two undergraduate assistants, who were representative of our subject pool.) A version of each of 64 pieces was generated for each condition, but each participant heard only one version of each piece, following a Latin square design. Each stimulus was a 24-s-long excerpt. For the Intact Music condition we used the original unmanipulated MIDI pieces. The Scrambled Music condition was produced via two manipulations of the MIDI files. First, a random number of semitones between -3 and 3 was added to the pitch of each note, to make the pitch distribution approximately uniform. The resulting pitch values were randomly reassigned to the notes of the piece, to remove contour structure. Second, to remove rhythmic structure, note onsets were jittered by a maximum of 1 beat (uniformly distributed), and note durations were randomly reassigned. The resulting piece had component sounds like those of the intact music but lacked high-level musical structure including key, rhythmic regularity, meter, and harmony. To examine potential dissociations between sensitivity to pitch and rhythmic scrambling, we also included two “intermediate” conditions: the Pitch Scrambled condition, in which only the note pitches were scrambled, and the Rhythm Scrambled condition, in which only the note onsets and durations were scrambled. Linear ramps (1 s) were applied to the beginning and end of each piece to avoid abrupt onsets/offsets. The scripts and sample stimuli are available at http://www.cns.nyu.edu/~jhm/music_scrambling/.

Our scrambling manipulation was intentionally designed to be relatively coarse. It has the advantage of destroying most of the melodic, harmonic, and rhythmic structure of music, arguably producing a more powerful contrast than has been used before. Given that previous scrambling manipulations have not revealed temporal lobe activations, it seemed important to use the strongest manipulation possible, which would be likely to reveal any brain regions sensitive to musical structure. However, the power of this contrast comes at the cost of some low-level differences between intact and scrambled conditions. We considered this trade-off to be worthwhile given our goal of probing temporal lobe sensitivity to music. We revisit this trade-off in DISCUSSION.

Stimuli were presented over scanner-safe earphones (Sensimetrics). At the beginning of the scan we ensured that the stimuli were clearly audible during a brief test run. For eight participants the task was to press a button after each piece, to help participants remain attentive. The last four participants were instead asked, “How much do you like this piece?” after each stimulus. Because the activation patterns were similar across the two tasks, we collapsed the data from these two subsets of participants. Condition order was counterbalanced across runs and participants. Experimental and fixation blocks lasted 24 and 16 s, respectively. Each run (16 experimental blocks—4 per condition—and 5 fixation blocks) lasted 464 s. Each participant completed four or five runs. Participants were instructed to avoid moving their fingers or feet in time with the music or humming/vocalizing with the music.

Language task. Participants read sentences, lists of unconnected words, and lists of unconnected pronounceable nonwords. In previous work we established that brain regions that are sensitive to high-level linguistic processing (defined by a stronger response to stimuli with syntactic and semantic structure, like sentences, than to meaningless and unstructured stimuli, like lists of nonwords) respond in a similar way to visually versus auditorily presented stimuli (Fedorenko et al. 2010; also Braze et al. 2011). We used visual presentation in the present study to ensure that the contrast between sentences (structured linguistic stimuli) and word lists (unstructured linguistic stimuli) reflected linguistic structure as opposed to possible prosodic differences (cf. Humphreys et al. 2005). Each stimulus consisted of eight words/nonwords. For details of how the language materials were constructed see Fedorenko et al. (2010). The materials are available at <http://web.mit.edu/evelina9/www/funcloc.html>.

Stimuli were presented in the center of the screen, one word/nonword at a time, at the rate of 350 ms per word/nonword. Each stimulus was followed by a 300-ms blank screen, a memory probe

¹ This sort of manipulation is analogous to those used to isolate structure processing in other domains. For example, contrasts between intact and scrambled pictures of objects have been used to study object processing (e.g., Malach et al. 1995). Similarly, contrasts between sentences and lists of unconnected words have been used to study syntactic and compositional semantic processing (e.g., Vandenberghe et al. 2002; Fedorenko et al. 2010).

² High-level linguistic structure can be contrasted with lower-level linguistic structure, like the sound structure of the language or the orthographic regularities for languages with writing systems.

(presented for 1,350 ms), and another blank screen for 350 ms, for a total trial duration of 4.8 s. Participants were asked to decide whether the probe appeared in the preceding stimulus by pressing one of two buttons. In previous work we established that similar brain regions are observed with passive reading (Fedorenko et al. 2010). Condition order was counterbalanced across runs and participants. Experimental and fixation blocks lasted 24 s (with 5 trials per block) and 16 s, respectively. Each run (12 experimental blocks—4 per condition—and 3 fixation blocks) lasted 336 s. Each participant completed four or five runs (with the exception of 1 participant who only completed 2 runs; because in our experience 2 runs are sufficient for eliciting robust language activations, this participant was included in all the analyses).

fMRI data acquisition. Structural and functional data were collected on the whole-body 3-T Siemens Trio scanner with a 32-channel head coil at the Athinoula A. Martinos Imaging Center at the McGovern Institute for Brain Research at MIT. T1-weighted structural images were collected in 128 axial slices with 1.33-mm isotropic voxels (TR = 2,000 ms, TE = 3.39 ms). Functional, blood oxygenation level-dependent (BOLD) data were acquired with an EPI sequence (with a 90° flip angle and using GRAPPA with an acceleration factor of 2), with the following acquisition parameters: thirty-one 4-mm-thick near-axial slices acquired in the interleaved order (with 10% distance factor), 2.1 mm × 2.1 mm in-plane resolution, FoV in the phase encoding (A >> P) direction 200 mm and matrix size 96 mm × 96 mm, TR = 2,000 ms, and TE = 30 ms. The first 10 s of each run were excluded to allow for steady-state magnetization.

fMRI data analyses. MRI data were analyzed with SPM5 (<http://www.fil.ion.ucl.ac.uk/spm>) and custom MATLAB scripts (available from <http://web.mit.edu/evelina9/www/funcloc>). Each subject's data were motion corrected and then normalized onto a common brain space [the Montreal Neurological Institute (MNI) template] and resampled into 2-mm isotropic voxels. Data were smoothed with a 4-mm Gaussian filter, high-pass filtered (at 200 s), and then analyzed in several different ways, as described next.

In the first analysis, to look for sensitivity to musical structure across the brain we conducted a whole-brain group-constrained subject-specific (GSS, formerly introduced as “GeSS”) analysis (Fedorenko et al. 2010; Julian et al. 2012). Because this analysis is relatively new, we provide a brief explanation of what it entails.

The goal of the whole-brain GSS analysis is to discover activations that are spatially similar across subjects without requiring voxel-level overlap (cf. the standard random-effects analysis; Holmes and Friston

1998), thus accommodating intersubject variability in the locations of functional activations (e.g., Frost and Goebel 2011; Pernet et al. 2007; Tahmasebi et al. 2012). Although the most advanced normalization methods (e.g., Fischl et al. 1999)—which attempt to align the folding patterns across individual brains—improve the alignment of functional activations compared with traditional methods, they are still limited because of the relatively poor alignment between cytoarchitecture (which we assume corresponds to function) and macroanatomy (sulci/gyri), especially in the lateral frontal and temporal cortices (e.g., Amunts et al. 1999; Brodmann 1909). The GSS method accommodates the variability across subjects in the locations of functional regions with respect to macroanatomy.

The GSS analysis includes the following steps: 1) Individual activation maps for the contrast of interest (i.e., Intact Music > Scrambled Music in this case) are thresholded (the threshold level will depend on how robust the activations are; we typically, including here, use the $P < 0.001$ uncorrected level) and overlaid on top of one another, resulting in a probabilistic overlap map, i.e., a map in which each voxel contains information on the percentage of subjects that show an above threshold response. 2) The probabilistic overlap map is divided into regions (“parcels”) by an image parcellation (watershed) algorithm. 3) The resulting parcels are then examined in terms of the proportion of subjects that show some suprathreshold voxels within their boundaries and the internal replicability.

The parcels that overlap with a substantial proportion of individual subjects and that show a significant effect in independent data (see below for the details of the cross-validation procedure) are considered meaningful. (For completeness, we include the results of the standard random-effects analysis in APPENDIX A.)

We focused on the parcels within which at least 8 of 12 individual subjects (i.e., ~67%; Fig. 1) showed suprathreshold voxels (at the $P < 0.001$ uncorrected level). However, to estimate the response of these regions to music and language conditions, we used the data from all 12 subjects, in order to be able to generalize the results in the broadest possible way,³ as follows. Each subject's activation map was computed for the Intact Music > Scrambled Music contrast using all

³ To clarify: if a functional region of interest (fROI) can only be defined in, e.g., 80% of the individual subjects, then the second-level results can be generalized to only 80% of the population (see Nieto-Castañón and Fedorenko 2012 for further discussion). Our method of defining fROIs in each subject avoids this problem. Another advantage of the approach whereby the top $n\%$ of the voxels within some anatomical/functional parcel are chosen in each individual is that the fROIs are identical in size across participants.

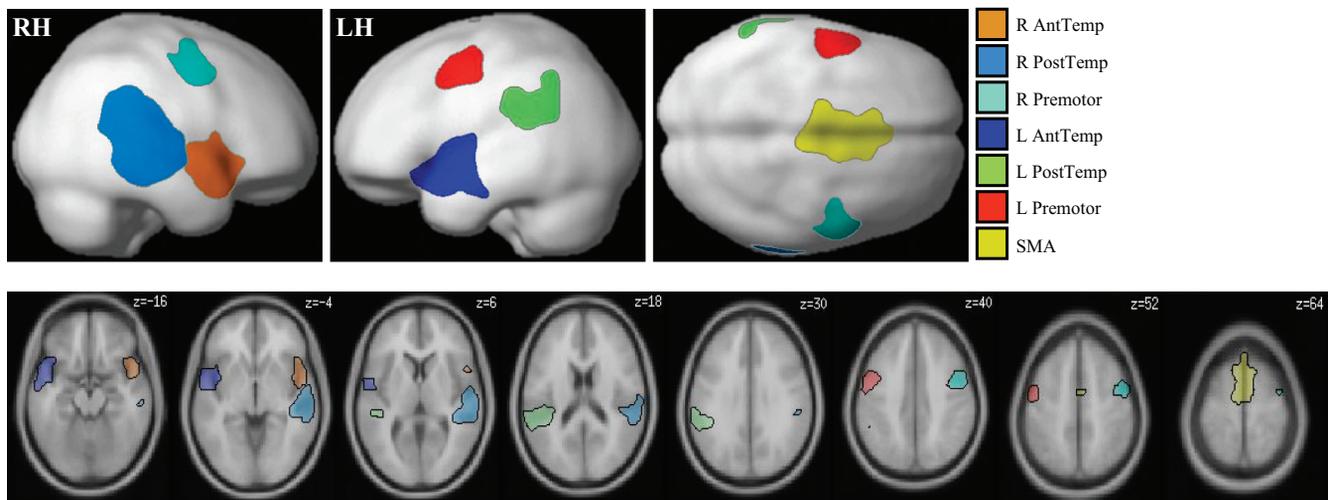


Fig. 1. *Top*: music-structure-sensitive parcels projected onto the surface of the brain. The parcels are regions within which most subjects (at least 8 of 12) showed above threshold activation for the Intact Music > Scrambled Music contrast ($P < 0.001$; see METHODS for details). *Bottom*: parcels projected onto axial slices (color assignment is similar to that used for the surface projection, with less saturated colors). For both surface and slice projection, we use the smoothed MNI template brain (avg152T1.nii template in SPM).

but one run of data, and the 10% of voxels with the highest t value within a given parcel (Fig. 1) were selected as that subject's fROI. The response was then estimated for this fROI using the left-out run. This procedure was iterated across all possible partitions of the data, and the responses were then averaged across the left-out runs to derive a single response magnitude for each condition in a given parcel/subject. This n -fold cross-validation procedure (where n is the number of functional runs) allows one to use all of the data for defining the ROIs and for estimating the responses (cf. the Neyman-Pearson lemma; see Nieto-Castañón and Fedorenko 2012 for further discussion), while ensuring the independence of the data used for fROI definition and for response estimation (Kriegeskorte et al. 2009).

Statistical tests across subjects were performed on the percent signal change (PSC) values extracted from the fROIs as defined above. Three contrasts were examined: 1) Intact Music > Scrambled Music to test for general sensitivity to musical structure; 2) Intact Music > Pitch Scrambled to test for sensitivity to pitch-related musical structure; and 3) Pitch Scrambled > Scrambled Music (both pitch and rhythm scrambled) to test for sensitivity to rhythm-related musical structure. The contrasts we used to examine sensitivity to pitch versus rhythm scrambling were motivated by an important asymmetry between pitch and timing information in music. Specifically, pitch information can be affected by the timing and order of different notes, while rhythm information can be appreciated even in the absence of pitch information (e.g., drumming). Consequently, to examine sensitivity to pitch scrambling, we chose to focus on stimuli with intact rhythmic structure, because scrambling the onsets of notes inevitably has a large effect on pitch-related information (for example, the grouping of different notes into chords). For the same reason, we used conditions whose pitch structure was scrambled to examine the effect of rhythm scrambling.

Because we observed sensitivity to the scrambling manipulation across extensive parts of the temporal lobes, we conducted a further GSS analysis to test whether there are lower-level regions that respond strongly to sounds but are insensitive to the scrambling of musical structure. To do so, we searched for voxels in each subject's brain that 1) responded more strongly to the Intact Music condition than to the baseline silence condition (at the $P < 0.001$, uncorrected, threshold) but that 2) did not respond more strongly to the Intact Music condition compared with the Scrambled Music condition ($P > 0.5$). Steps 1, 2, and 3 of the GSS analysis were then performed as described above. Also as in the above analysis, we focused on parcels within which at least 8 of 12 individual subjects (i.e., ~67%) showed voxels with the specified functional properties.

In the second analysis, to examine the responses of the music-structure-sensitive fROIs to high-level linguistic structure, we used the same fROIs as in the first analysis and extracted the PSC values for the Sentences and Word Lists conditions. Statistical tests were performed on these values. The contrast Sentences > Word Lists was examined to test for sensitivity to high-level linguistic structure (i.e., syntactic and/or compositional semantic structure).

To demonstrate that the Sentences > Word Lists contrast engages regions that have been previously identified as sensitive to linguistic structure (Fedorenko et al. 2010), we also report the response profiles of brain regions sensitive to high-level linguistic processing, defined by the Sentences > Nonword Lists contrast. We report the responses of these regions to the three language conditions (Sentences, Word Lists, and Nonword Lists); the responses to the Sentences and Nonword Lists conditions are estimated with cross-validation across runs) and to the Intact Music and Scrambled Music conditions. These data are the same as those reported previously by Fedorenko et al. (2011), except that the fROIs are defined by the top 10% of the Sentences > Nonword Lists voxels, rather than by the hard threshold of $P < 0.001$, uncorrected. This change was made to make the analysis consistent with the other analyses in this report; the results are similar regardless of the details of the fROI definition procedure.

RESULTS

Looking for sensitivity to musical structure across the brain. The GSS analysis revealed seven parcels (Fig. 1) in which the majority of subjects showed a greater response to intact than scrambled music. In the remainder of this article we will refer to these regions as “music-structure-sensitive” regions. These include bilateral parcels in the anterior superior temporal gyrus (STG) (anterior to the primary auditory cortex), bilateral parcels in the posterior STG (with the right hemisphere parcel also spanning the middle temporal gyrus⁴), bilateral parcels in the premotor cortex, and the supplementary motor area (SMA). Each of the seven regions showed a significant effect for the Intact Music > Scrambled Music contrast, estimated with independent data from all 12 subjects in the experiment ($P < 0.01$ in all cases; Table 1).

Our stimulus scrambling procedure allowed us to separately examine the effects of pitch and rhythm scrambling. In Fig. 2 we present the responses of our music-structure-sensitive fROIs to all four conditions of the music experiment (estimated with cross-validation, as described in METHODS). In each of these regions we found significant sensitivity to both the pitch scrambling and rhythm scrambling manipulations (all $P < 0.05$; Table 1).

One could argue that it is unsurprising that the responses to the Pitch Scrambled and Rhythm Scrambled conditions fall in between the Intact Music and the Scrambled Music conditions given that the Intact Music > Scrambled Music condition was used to localize the regions. It is worth noting that this did not have to be the case: for example, some regions could show the Intact Music > Scrambled Music effect because the Intact Music condition has a pitch contour; in that case, the Rhythm Scrambled condition—in which the pitch contour is preserved—might be expected to pattern with the Intact Music condition, and the Pitch Scrambled condition with the Scrambled Music condition. Nevertheless, to search for regions outside of those that respond more to intact than scrambled music, as well as for potential subregions within the music-structure-sensitive regions, we performed additional whole-brain GSS analyses on the narrower contrasts (i.e., Pitch Scrambled > Scrambled Music and Rhythm Scrambled > Scrambled Music). If some regions outside of the borders of our Intact Music > Scrambled Music regions, or within their boundaries, are selectively sensitive to pitch contour or rhythmic structure, the GSS analysis on these contrasts should discover those regions. Because these contrasts are functionally narrower and because we wanted to make sure not to miss any regions, we tried these analyses with thresholding individual maps at both $P < 0.001$ (as for the Intact Music > Scrambled Music contrast reported here) and a more liberal, $P < 0.01$ level. The regions that emerged for these contrasts 1) fell within the broader Intact Music > Scrambled Music regions and 2) showed response profiles similar to those of the Intact Music > Scrambled Music regions, suggesting that we

⁴ Because we were concerned that the RPostTemp parcel was large, spanning multiple anatomical structures, we performed an additional analysis in which prior to its parcellation the probabilistic overlap map was thresholded to include only voxels where at least a quarter of the subjects (i.e., at least 3 of the 12) showed the Intact Music > Scrambled Music effect (at the $P < 0.001$ level or higher). The resulting much smaller parcel—falling largely within the middle temporal gyrus—showed the same functional properties as the original parcel (see APPENDIX B).

Table 1. Effects of music scrambling manipulations in regions discovered by GSS analysis

	Right Hemisphere		Left Hemisphere		SMA		
	AntTemp	PostTemp	Premotor	AntTemp		PostTemp	Premotor
Step 1. Sensitivity to music structure (IM>SM)	$t = 3.83; P < 0.005$	$t = 3.09; P < 0.01$	$t = 2.96; P < 0.01$	$t = 3.45; P < 0.005$	$t = 4.54; P < 0.001$	$t = 4.33; P < 0.001$	$t = 3.30; P < 0.005$
Step 2. Sensitivity to pitch scrambling (IM>PS)	$t = 3.14; P < 0.005$	$t = 2.30; P < 0.05$	$t = 2.40; P < 0.05$	$t = 4.18; P < 0.001$	$t = 4.39; P < 0.001$	$t = 3.55; P < 0.005$	$t = 3.18; P < 0.005$
Step 3. Sensitivity to rhythm scrambling (PS>SM)	$t = 3.43; P < 0.005$	$t = 3.86; P < 0.005$	$t = 2.90; P < 0.01$	$t = 2.55; P < 0.05$	$t = 3.77; P < 0.005$	$t = 3.60; P < 0.005$	$t = 2.90; P < 0.01$

GSS, group-constrained subject-specific; SMA, supplementary motor area; IM, Intact Music; SM, Scrambled Music; PS, Pitch Scrambled. We report uncorrected P values ($df = 11$), but all effects remain significant after an FDR correction for the number of regions ($n = 7$).

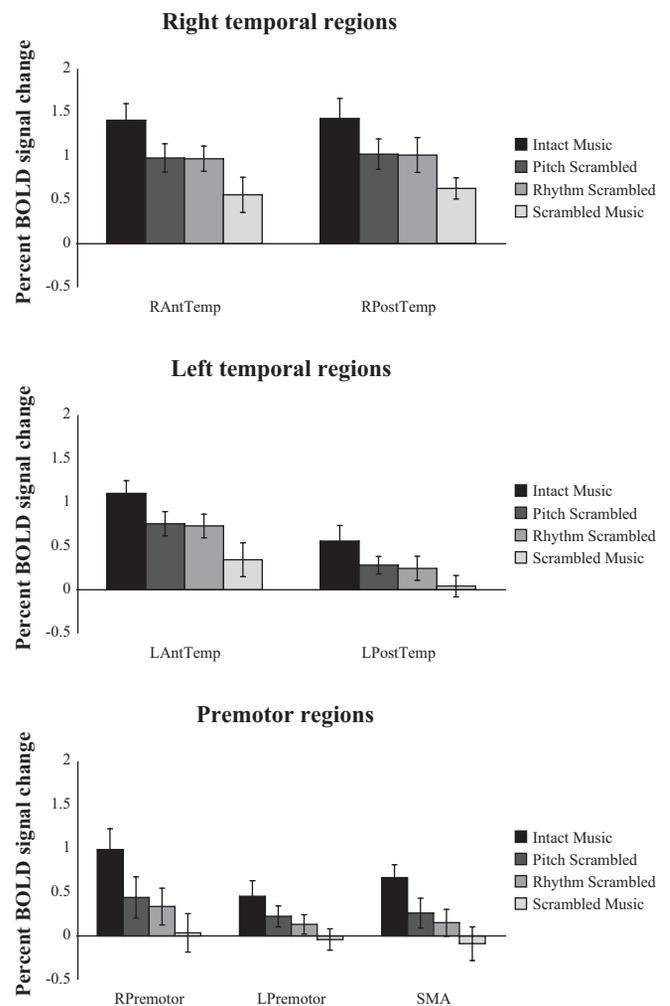


Fig. 2. Responses of music-structure-sensitive regions discovered by the group-constrained subject-specific (GSS) analysis and defined in each individual participant with the 10% of voxels in a given parcel with the most significant response to the Intact Music > Scrambled Music contrast. The responses are estimated by n -fold cross-validation, as discussed in METHODS, so that the data used to define the functional regions of interest (fROIs) and estimate the responses are independent. Error bars reflect SE. BOLD, blood oxygenation level dependent.

are not missing any regions selectively sensitive to the pitch contour or rhythmic structure.

The “control” GSS analysis revealed three parcels (Fig. 3) that responded strongly to all four music conditions but showed no sensitivity to the scrambling manipulation (replicating the search criteria in independent data). These parcels fell in the posterior portion of the STG/superior temporal sulcus (STS), overlapping also with Heschl’s gyrus, and thus plausibly corresponding to primary auditory regions. Each of the three regions showed a significant effect for the Intact Music > Baseline contrast, estimated in independent data (all $t > 6$, all $P < 0.0005$) but no difference between the Intact and Scrambled Music conditions (all $t < 1.1$, not significant; Fig. 4). [Note that although these parcels may spatially overlap with the music-structure-sensitive parcels discussed above, individual fROIs are defined by intersecting the parcels with each subject’s activation map. As a result, the music-structure-sensitive and control fROIs are unlikely to overlap in individual subjects.]

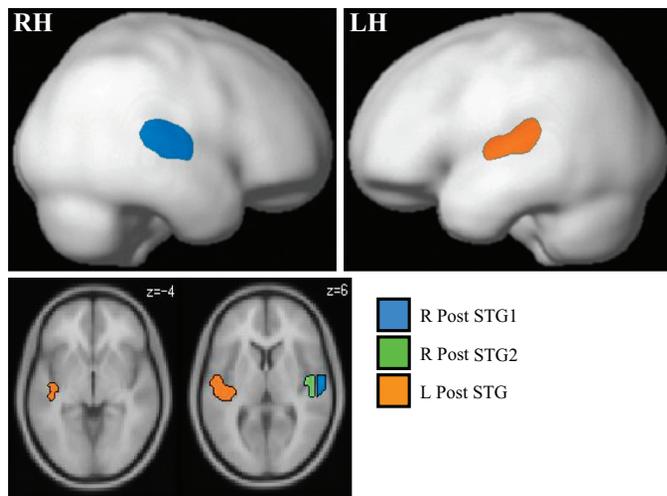


Fig. 3. *Top*: parcels from the “control” GSS analysis projected onto the surface of the brain (only 2 of the 3 parcels are visible on the surface). The parcels are regions within which most subjects (at least 8 of 12) showed voxels that responded robustly to Intact Music but did not differ in their response to Intact vs. Scrambled Music conditions (see METHODS for details). *Bottom*: parcels projected onto axial slices (color assignment is similar to that used for the surface projection). For both surface and slice projection we use the smoothed MNI template brain (avg152T1.nii template in SPM).

Sensitivity to musical structure vs. high-level linguistic structure. In Fig. 5, *top*, we show the responses of Intact Music > Scrambled Music fROIs to the three conditions of the language experiment (Sentences, Word Lists, and Nonword Lists). Although the music-structure-sensitive regions respond above baseline to the language conditions, none shows sensitivity to linguistic structure, responding similarly to the Sentences and Word Lists conditions (all $t < 1$).

In Fig. 5, *bottom*, we show the responses of brain regions sensitive to high-level linguistic structure (defined as responding more strongly to the Sentences condition than to the Nonword Lists condition) to the language and music conditions. The effect of linguistic structure (Sentences > Word Lists) was robust in all of the language fROIs (all $t > 3.4$, all $P < 0.005$). These effects demonstrate that the lack of sensitivity to high-level linguistic structure in the music fROIs is not due to the ineffectiveness of the manipulation: the Sentences > Word Lists contrast activates extended portions of the left frontal and temporal cortices (see Fedorenko and Kanwisher

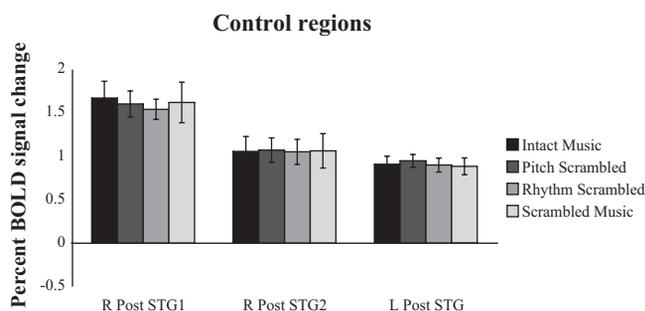


Fig. 4. Responses of the control regions discovered by the GSS analysis that respond to the Intact Music condition but do not show the Intact Music > Scrambled Music effect. The ROIs are defined functionally with a conjunction of the Intact Music > Baseline contrast and a negation of the Intact Music > Scrambled Music contrast. The responses are estimated by n -fold cross-validation, as discussed in METHODS, so that the data used to define the fROIs and estimate the responses are independent.

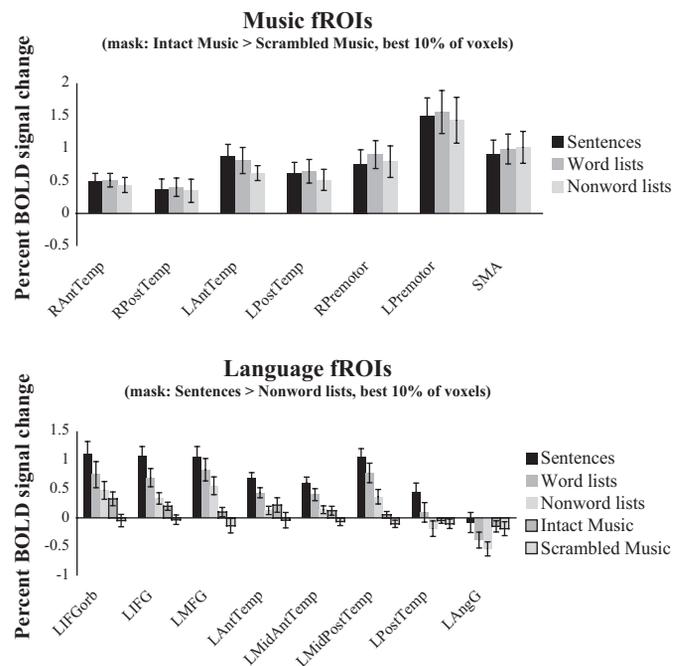


Fig. 5. Double dissociation: music fROIs are not sensitive to linguistic structure, and language fROIs are not sensitive to music structure. *Top*: responses of music-structure-sensitive regions to the language conditions (regions were defined in each individual participant with the 10% of voxels within each parcel that had the highest t values for the Intact Music > Scrambled Music comparison, as described in METHODS). *Bottom*: responses of brain regions sensitive to high-level linguistic processing to the language and music conditions [regions were defined in each individual participant with the 10% of voxels within each parcel that had the highest t values for the Sentences > Nonword Lists comparison; parcels were taken from Fedorenko et al. (2010)]. [With the exception of the Word Lists condition, these data were reported in Fedorenko et al. (2011).]

2011 for sample individual whole-brain activation maps for this contrast). However, although several of the language fROIs show a stronger response to the Intact Music than the Scrambled Music condition (with a few regions reaching significance at the $P < 0.05$ uncorrected level: LIFGorb, LIFG, LAntTemp, LMidAntTemp, and LMidPostTemp), this effect does not survive the FDR correction for the number of regions ($n = 8$). Additionally, in only two of the regions (LIFGorb and LIFG) is the response to the Intact Music condition reliably greater than the response to the fixation baseline condition⁵ (compare to the temporal musical-structure-sensitive regions, in which this difference is highly robust: $P < 0.0001$ in the right AntTemp and PostTemp regions and in the left AntTemp region; $P < 0.005$ in the left PostTemp region). The overall low response to intact music suggests that these regions are less relevant to the processing of musical structure than are the temporal regions we found to be sensitive to music scrambling.

DISCUSSION

Our results revealed several brain regions that showed apparent sensitivity to music structure, as evidenced by a stronger response to intact than scrambled musical stimuli. These regions include

⁵ Note that the lack of a large response to music relative to the fixation baseline in the language fROIs is not because these regions only respond to visually presented stimuli. For example, in Fedorenko et al. (2010) we report robust responses to auditorily presented linguistic stimuli in these same regions.

anterior parts of the STG bilaterally and posterior parts of the superior and middle temporal gyri bilaterally, as well as premotor regions and the SMA. A control analysis revealed brain regions in and around primary auditory cortices that robustly responded to intact musical stimuli—similar to the regions above—and yet showed no difference between intact and scrambled musical stimuli, in contrast to regions sensitive to musical structure. The latter result suggests that sensitivity to musical structure is mainly limited to regions outside of primary auditory cortex. We draw three main conclusions from our findings. First, and most importantly, sensitivity to musical structure is robustly present in the temporal lobes, consistent with the patient literature. Second, each of the music-structure-sensitive brain regions shows sensitivity to both pitch and rhythm scrambling. And third, there exist brain regions that are sensitive to musical but not high-level linguistic structure, again as predicted by patient findings (Luria et al. 1965; Peretz and Coltheart 2003).

Brain regions sensitive to musical structure. Previous patient and neuroimaging studies have implicated brain regions anterior and posterior to primary auditory cortex in music processing, but their precise contribution to music remains an open question (for reviews see, e.g., Griffiths and Warren 2002; Koelsch 2011; Koelsch and Siebel 2005; Limb 2006; Patel 2003, 2008; Peretz and Zatorre 2005; Samson et al. 2011; Zatorre and Schoenwiesner 2011). In the present study we found that regions anterior and posterior to Heschl's gyrus in the superior temporal plane (PP and PT) as well as parts of the superior and middle temporal gyri respond more to intact than scrambled musical stimuli, suggesting a role in the analysis or representation of musical structure.

Why haven't previous neuroimaging studies that used scrambling manipulations observed sensitivity to musical structure in the temporal lobe? A likely reason is that our manipulation scrambles musically relevant structure more drastically than previous manipulations. In particular, previous scrambling procedures have largely preserved local musical structure (e.g., by rearranging ~300-ms-long chunks of music; Levitin and Menon 2003), to which temporal regions may be sensitive. There is, of course, also a cost associated with the use of a relatively coarse manipulation of musical structure: the observed responses could in part be driven by factors unrelated to music (e.g., lower-level pitch and timing differences; e.g., Zatorre and Belin 2001). Reassuringly though, bilateral regions in the posterior STG/Heschl's gyrus, in and around primary auditory cortex, showed similarly strong responses to intact and scrambled musical stimuli. Thus, although it is difficult to rule out the contribution of low-level differences to the scrambling effects we observed, we think it is likely that the greater response to intact than scrambled music stimuli is at least partly due to the presence of (Western) musical structure (e.g., key, meter, harmony, melodic contour), particularly in the higher-order temporal regions.

What is the function of the music-structure-sensitive brain regions? One possibility is that these regions store musical knowledge⁶ [what Peretz and Coltheart (2003) refer to as the "musical lexicon"], which could include information about

melodic and/or rhythmic patterns that are generally likely to occur (presumably learned from exposure to music), as well as memories of specific musical sequences ("musical schemata" and "musical memories", respectively; Justus and Bharucha 2001; also Bharucha and Stoeckig 1986; Patel 2003; Tillmann et al. 2000). The response in these regions could therefore be a function of how well the stimulus matches stored representations of prototypical musical structures.

It is also possible that some of the observed responses reflect sensitivity to more generic types of structure in music. For example, the scrambling procedure used here affects the overall consonance/dissonance of simultaneous and adjacent notes, which may be important given that pitch-related responses have been reported in anterior temporal regions similar to those observed here (Norman-Haignere et al. 2011; Patterson et al. 2002; Penagos et al. 2004) and given that consonance perception appears to be closely related to pitch processing (McDermott et al. 2010; Terhardt 1984). In addition, the scrambling procedure affects the distribution and variability of interonset note intervals as well as the coherence of different musical streams/melodic lines. Teasing apart sensitivity to generic versus music-specific structure will be an important goal for future research.

In addition to the temporal lobe regions, we also found sensitivity to music scrambling in bilateral premotor regions and in the SMA. These regions are believed to be important for planning complex movements and have been reported in several neuroimaging studies of music, including studies of musicians listening to pieces they can play, which presumably evokes motor imagery (e.g., Bangert et al. 2006; Baumann et al. 2005), as well as studies on beat perception and synchronization (e.g., Chen et al. 2006; Grahn and Brett 2007; Kornysheva et al. 2010). Although one might have predicted that rhythm structure would be more important than melodic structure for motor areas, pitch and rhythmic structure are highly interdependent in music (e.g., Jones and Boltz 1989), and thus scrambling pitch structure may have also affected the perceived rhythm/meter.

Sensitivity to pitch vs. rhythm scrambling. Musical pitch and rhythm are often separated in theoretical discussions (e.g., Krumhansl 2000; Lerdahl and Jackendoff 1983). Furthermore, some evidence from amusic patients and neuroimaging studies suggests that mechanisms that support musical pitch and rhythmic processing may be distinct, with some studies further suggesting that the right hemisphere may be especially important for pitch perception and the left hemisphere more important for rhythm perception (see, e.g., Peretz and Zatorre 2005 for a summary). However, we found that each of the brain regions that responded more to intact than scrambled music showed sensitivity to both pitch and rhythm scrambling manipulations (see also Griffiths et al. 1999). This surprising result may indicate that the processing of pitch and rhythm are inextricably linked (e.g., Jones and Boltz 1989), a conclusion that would have important implications for our ultimate understanding of the cognitive and neural mechanisms underlying music. In an intriguing parallel, current evidence suggests a similar overlap in brain regions sensitive to lexical meanings and syntactic/compositional semantic structure in language (e.g., Fedorenko et al. 2012). It is worth noting, however, that even though the responses of all the music-structure-sensitive regions were affected by both pitch and rhythm scrambling, these regions may differ with respect to their causal role in

⁶ One could hypothesize that musical memories are instead stored in the hippocampus and adjacent medial temporal lobe structures, which are implicated in the storage of episodic memories. However, Finke et al. (2012) recently provided evidence against this hypothesis, by demonstrating that a professional cellist who developed severe amnesia following encephalitis nevertheless performed similarly to healthy musicians on tests of music recognition.

processing pitch versus rhythm, as could be probed with transcranial magnetic stimulation in future work.

Sensitivity to musical vs. high-level linguistic structure. None of the regions that responded more to intact than scrambled musical stimuli showed sensitivity to high-level linguistic structure (i.e., to the presence of syntactic and semantic relationships among words), suggesting that it is not the case that these regions respond more to any kind of structured compared with unstructured/scrambled stimulus. This lack of sensitivity to linguistic structure in the music-structure-sensitive regions is notable given that language stimuli robustly activate extended portions of the frontal and temporal lobes, especially in the left hemisphere (e.g., Binder et al. 1997; Fedorenko et al. 2010; Neville et al. 1998). However, these results are consistent with two recent reports of the lack of sensitivity to musical structure in brain regions that are sensitive to high-level linguistic structure (Fedorenko et al. 2011; Rogalsky et al. 2011). For completeness, we report data from Fedorenko et al. (2011) (which used the same linguistic stimuli that we used to probe our music parcels) in the present article. Brain regions sensitive to high-level linguistic processing showed robust sensitivity to linguistic structure (in independent data), responding significantly more strongly to the Sentences condition, which involves syntactic and compositional semantic structure, than to the Word Lists condition, which has neither syntactic nor compositional semantic structure. However, the response to the Intact Music condition in these regions was low, even though a few ROIs (e.g., LIFGorb) showed a somewhat higher response to intact than scrambled stimuli, consistent with Levitin and Menon (2003). Although this sensitivity could be functionally important, possibly consistent with the “neural re-use” hypotheses (e.g., Anderson 2010), these effects should be interpreted in the context of the overall much stronger response to linguistic than musical stimuli.

The existence of the regions identified here that respond to musical structure but not linguistic structure does not preclude the existence of other regions that may in some way be engaged by the processing of both musical and linguistic stimuli (e.g., Francois and Schon 2011; Janata and Grafton 2003; Koelsch et al. 2002; Maess et al. 2001; Merrill et al. 2012; Osnes et al. 2012; Patel 2003; Tillmann et al. 2003). As noted in the introduction, these previously reported regions of overlap appear to be engaged in a wide range of demanding cognitive tasks, including those that have little to do with music or hierarchical structural processing (e.g., Corbetta and Shulman 2002; Duncan 2001, 2010; Duncan and Owen 2000; Miller and Cohen 2001). Consistent with the idea that musical processing engages some domain-general mechanisms, several studies have now shown that musical training leads to improvement in general executive functions, such as working memory and attention (e.g., Besson et al. 2011; Moreno et al. 2011; Neville et al. 2009; Sluming et al. 2007; Strait and Kraus 2011; cf. Schellenberg 2011). Similarly, our findings are orthogonal to the question of whether overlap exists in the lower-level acoustic processes in music and speech (e.g., phonological or prosodic processing). Indeed, previous research has suggested that pitch processing in speech and music may rely on shared encoding mechanisms in the auditory brain stem (Krizman et al. 2012; Parbery-Clark et al. 2012; Strait et al. 2011; Wong et al. 2007).

Conclusions. Consistent with findings from the patient literature, we report several regions in the temporal cortices that are sensitive to musical structure and yet show no response to high-level linguistic (syntactic/compositional semantic) structure. These regions are candidates for the neural basis of music. The lack of sensitivity of these regions to high-level linguistic structure suggests that the uniquely and universally human capacity for music is not based on the same mechanisms as our species’ other famously unique capacity for language. Future work can now target these candidate “music regions” to examine neural specialization for music and to characterize the representations they store and the computations they perform.

APPENDIX A

Results of Traditional Random-Effects Analysis for Intact Music > Scrambled Music Contrast

In Fig. 6 we show the results of the traditional random-effects group analysis for the Intact Music > Scrambled Music contrast. This analysis reveals several clusters of activated voxels, including 1) bilateral clusters in the STG anterior to primary auditory cortex (in the planum polare), 2) a small cluster in the right posterior temporal lobe that falls mostly within the middle temporal gyrus, and 3) several clusters in the right frontal lobe, including both right IFG, consistent with Levitin and Menon’s (2003) findings, and right middle frontal gyrus (see Table 2).

APPENDIX B

Additional Analysis for RPostTemp Parcel

Because the parcel that was discovered in the original GSS analysis in the right posterior temporal cortex was quite large, spanning multiple anatomical structures, we performed an additional analysis in which prior to its parcellation the probabilistic overlap map was thresholded to include only voxels where at least a quarter of the subjects (i.e., at least 3 of the 12) showed the Intact Music > Scrambled Music effect (at the $P < 0.001$ level or higher). Such thresholding has two consequences: 1) parcels decrease in size and 2) fewer subjects may show suprathreshold voxels within the parcel. In Fig. 7, *left*, we show the original RPostTemp parcel (in turquoise) and the parcel that resulted from the new analysis (in green). The new parcel falls largely within the middle temporal gyrus. Nine of the twelve subjects showed voxels within the boundaries of the new parcel that reached significance at the $P < 0.001$ level at the whole-brain level.

To estimate the response profile of this region, we used the same procedure as in the analysis reported above. In particular, we used the 10% of voxels with the highest Intact Music > Scrambled Music voxels in each subject within the parcel for all but the first run of the data. We then iteratively repeated the procedure across all possible partitions of the data and averaged the responses across the left-out

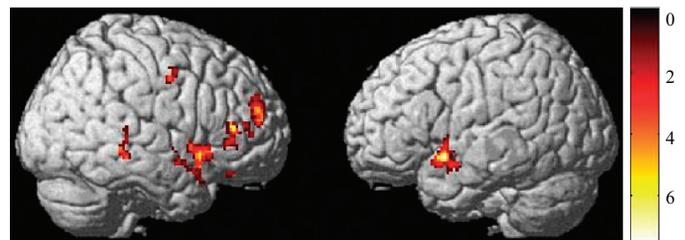


Fig. 6. Activation map from the random effects analysis for the Intact Music > Scrambled Music contrast (thresholded at $P < 0.001$, uncorrected) projected onto the single-subject template brain in SPM (single_subj_T1.img).

Table 2. Random-effects analysis results for Intact Music > Scrambled Music contrast

Cluster Level			Voxel Level					x, y, z, mm
P corr.	k _E	P uncorr.	P FWEcorr.	P FDRcorr.	T	Z	P uncorr.	
0.000	297	0.000	0.012	0.012	11.66	5.25	<0.0001	40 2 -14
			0.038	0.018	10.41	5.03	<0.0001	46 14 -8
			0.903	0.076	7.09	4.26	<0.0001	32 2 -20
0.317	29	0.008	0.515	0.051	7.98	4.50	<0.0001	22 34 18
0.617	21	0.021	0.896	0.076	7.12	4.27	<0.0001	18 34 -20
0.055	48	0.001	0.913	0.076	7.04	4.25	<0.0001	-20 -62 -28
			1.000	0.100	5.33	3.67	<0.0001	-26 -62 -18
			1.000	0.100	4.74	3.43	<0.0001	-20 -56 -18
0.617	21	0.021	0.944	0.089	6.85	4.19	<0.0001	8 58 10
0.000	180	0.000	0.963	0.090	6.71	4.15	<0.0001	-42 8 -10
			1.000	0.100	5.80	3.85	<0.0001	-42 2 -16
			1.000	0.100	5.48	3.73	<0.0001	-54 6 -10
0.003	83	0.000	0.963	0.090	6.70	4.15	<0.0001	40 34 8
			1.000	0.100	5.39	3.69	<0.0001	42 36 0
			1.000	0.100	4.71	3.41	<0.0001	50 40 6
0.000	156	0.000	1.000	0.100	5.71	3.82	<0.0001	28 52 12
			1.000	0.100	5.62	3.78	<0.0001	22 52 20
			1.000	0.100	4.53	3.33	<0.0001	20 44 16
0.530	23	0.017	1.000	0.100	5.40	3.70	<0.0001	6 6 24
0.289	30	0.008	1.000	0.100	5.39	3.69	<0.0001	48 -8 42
			1.000	0.100	4.43	3.29	<0.0001	54 -2 46
0.066	46	0.002	1.000	0.100	5.22	3.63	<0.0001	58 -38 -6
			1.000	0.100	4.62	3.37	<0.0001	62 -34 8
0.573	22	0.019	1.000	0.100	4.64	3.38	<0.0001	34 44 10

Height threshold $T = 4.025$ ($P < 0.001$, uncorrected). Extent threshold $k = 20$ voxels. Significance values are adjusted for search volume.

runs. The responses of the smaller RPostTemp fROI to the four music conditions are shown in Fig. 7, right. As expected, the responses are similar to those observed for the original fROI because we are simply narrowing in on the peak part of the original parcel. The statistics for the three contrasts examining sensitivity to musical scrambling were as follows: general sensitivity to musical structure: Intact Music > Scrambled Music, $t(11) = 3.49$, $P < 0.005$; sensitivity to pitch scrambling: Intact Music > Pitch Scrambled, $t(11) = 3.35$, $P < 0.005$; sensitivity to rhythm scrambling: Pitch Scrambled > Scrambled Music, $t(11) = 2.75$, $P < 0.01$.

Furthermore, as in the original analysis, the new RPostTemp fROI showed no sensitivity to linguistic structure, responding similarly strongly to lists of words and sentences: 0.40 (SE = 0.25) and 0.44 (SE = 0.23), respectively ($t < 1$).

ACKNOWLEDGMENTS

We thank Tanya Goldhaber for help in creating the music stimuli, Jason Webster, Eyal Dechter, and Michael Behr for help with the experimental scripts and with running the participants, and Christina Triantafyllou, Steve

Shannon, and Sheeba Arnold for technical support. We thank the members of the Kanwisher, Gibson, and Saxe labs for helpful discussions and the audiences at the Neurobiology of Language conference in 2011 and at the CUNY sentence processing conference in 2012 for helpful comments. We also acknowledge the Athinoula A. Martinos Imaging Center at McGovern Institute for Brain Research, MIT.

GRANTS

This research was supported by Eunice Kennedy Shriver National Institute of Child Health and Human Development Award K99 HD-057522 to E. Fedorenko and a grant to N. Kanwisher from the Ellison Medical Foundation. J. H. McDermott was supported by the Howard Hughes Medical Institute.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

Author contributions: E.F., J.H.M., and N.K. conception and design of research; E.F. and S.N.-H. performed experiments; E.F. analyzed data; E.F.,

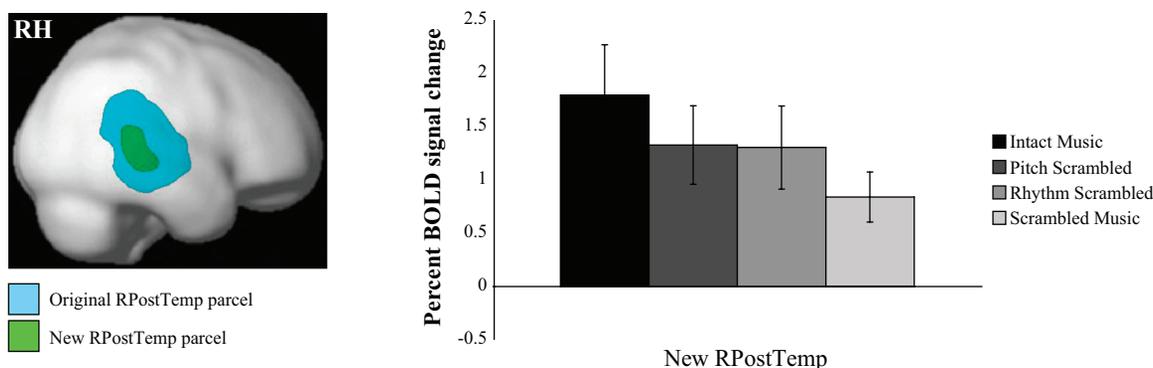


Fig. 7. Left: original and new RPostTemp parcels projected onto the lateral surface. Right: responses of the new fROI to the 4 music conditions (estimated with cross-validation, as in the analyses in the text).

J.H.M., S.N.-H., and N.K. interpreted results of experiments; E.F. prepared figures; E.F. drafted manuscript; E.F., J.H.M., S.N.-H., and N.K. edited and revised manuscript; E.F., J.H.M., S.N.-H., and N.K. approved final version of manuscript.

REFERENCES

- Abrams D, Bhatara A, Ryali S, Balaban E, Levitin DJ, Menon V.** Decoding temporal structure in music and speech relies on shared brain resources but elicits different fine-scale spatial patterns. *Cereb Cortex* 21: 1507–1518, 2011.
- Allen G.** Note-deafness. *Mind* 10: 157–167, 1878.
- Amunts K, Schleicher A, Burgel U, Mohlberg H, Uylings HB, Zilles K.** Broca's region revisited: cytoarchitecture and inter-subject variability. *J Comp Neurol* 412: 319–341, 1999.
- Anderson ML.** Neural re-use: a fundamental organizational principle of the brain. *Behav Brain Sci* 33: 245–266, 2010.
- Bangert M, Peschel T, Schlaug G, Rotte M, Drescher D, Hinrichs H, Heinze HJ, Altenmüller E.** Shared networks for auditory and motor processing in professional pianists: evidence from fMRI conjunction. *Neuroimage* 30: 917–926, 2006.
- Baumann S, Koenke S, Meyer M, Lutz K, Jancke L.** A network for sensory-motor integration: what happens in the auditory cortex during piano playing without acoustic feedback? *Ann NY Acad Sci* 1060: 186–188, 2005.
- Besson M, Faita F.** An event-related potential (ERP) study of musical expectancy: comparison of musicians with nonmusicians. *J Exp Psychol Hum Percept Perform* 21: 1278–1296, 1995.
- Besson M, Chobert J, Marie C.** Transfer of training between music and speech: common processing, attention, and memory. *Front Psychol* 2: 94, 2011.
- Bharucha JJ.** Anchoring effects in music: the resolution of dissonance. *Cogn Psychol* 16: 485–518, 1984.
- Bharucha JJ, Stoeckig K.** Reaction time and musical expectancy: priming of chords. *J Exp Psychol Hum Percept Perform* 12: 403–410, 1986.
- Binder JR, Frost JA, Hammeke TA, Bellgowan P, Springer JA, Kaufman JN, Possing ET.** Human temporal lobe activation by speech and non-speech sounds. *Cereb Cortex* 10: 512–520, 2000.
- Binder JR, Frost JA, Hammeke TA, Cox RW, Rao SM, Thomas P.** Human brain language areas identified by functional magnetic resonance imaging. *J Neurosci* 17: 353–362, 1997.
- Braze D, Mencl WE, Tabor W, Pugh KR, Constable RT, Fulbright RK, Magnuson JS, Van Dyke JA, Shankweiler DP.** Unification of sentence processing via ear and eye: an fMRI study. *Cortex* 47: 416–431, 2011.
- Brodman K.** *Vergleichende Lokalisationslehre der Grosshirnrinde in Ihren Prinzipien dargestellt auf Grund des Zellenbaues.* Leipzig, Germany: Barth, 1909.
- Chen JL, Zatorre R, Penhune V.** Interactions between auditory and dorsal premotor cortex during synchronization to musical rhythms. *Neuroimage* 32: 1771–1781, 2006.
- Corbetta M, Shulman GL.** Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci* 3: 215–229, 2002.
- Dalla Bella S, Peretz I.** Music agnosias: selective impairments of music recognition after brain damage. *J New Music Res* 28: 209–216, 1999.
- Di Pietro M, Laganaro M, Leemann B, Schnider A.** Receptive amusia: temporal auditory processing deficit in a professional musician following a left temporo-parietal lesion. *Neuropsychologia* 42: 868–877, 2004.
- Doeller C, Opitz B, Mecklinger A, Krick C, Reith W, Schroger E.** Prefrontal cortex involvement in preattentive auditory deviance detection: neuroimaging and electrophysiological evidence. *Neuroimage* 20: 1270–1282, 2003.
- Dowling WJ, Harwood DL.** *Music Cognition.* Orlando, FL: Academic, 1986.
- Duncan J.** An adaptive coding model of neural function in prefrontal cortex. *Nat Rev Neurosci* 2: 820–829, 2001.
- Duncan J.** The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. *Trends Cogn Sci* 14: 172–179, 2010.
- Duncan J, Owen AM.** Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends Neurosci* 23: 475–483, 2000.
- Evers S, Dannert J, Rodding D, Rotter G, Ringelstein EB.** The cerebral haemodynamics of music perception: a transcranial Doppler sonography study. *Brain* 122, 75–85, 1999.
- Fedorenko E, Behr M, Kanwisher N.** Functional specificity for high-level linguistic processing in the human brain. *Proc Natl Acad Sci USA* 108: 16428–16433, 2011.
- Fedorenko E, Hsieh PJ, Nieto-Castañón A, Whitfield-Gabrieli S, Kanwisher N.** A new method for fMRI investigations of language: defining ROIs functionally in individual subjects. *J Neurophysiol* 104: 1177–1194, 2010.
- Fedorenko E, Kanwisher N.** Some regions within Broca's area do respond more strongly to sentences than to linguistically degraded stimuli: a comment on Rogalsky and Hickok (2010). *J Cogn Neurosci* 23: 2632–2635, 2011.
- Fedorenko E, Nieto-Castañón A, Kanwisher N.** Lexical and syntactic representations in the brain: an fMRI investigation with multi-voxel pattern analyses. *Neuropsychologia* 50: 499–513, 2012.
- Fedorenko E, Patel A, Casasanto D, Winawer J, Gibson E.** Structural integration in language and music: evidence for a shared system. *Mem Cognit* 37: 1–9, 2009.
- Finke C, Esfahani NE, Ploner CJ.** Preservation of musical memory in an amnesic professional cellist. *Curr Biol* 22: R591–R592, 2012.
- Fischl B, Sereno MI, Tootell RB, Dale AM.** High-resolution intersubject averaging and a coordinate system for the cortical surface. *Hum Brain Mapp* 8: 272–284, 1999.
- Francois C, Schon D.** Musical expertise boosts implicit learning of both musical and linguistic structures. *Cereb Cortex* 21: 2357–2365, 2011.
- Frost MA, Goebel R.** Measuring structural-functional correspondence: spatial variability of specialized brain regions after macro-anatomical alignment. *Neuroimage* 59: 1369–1381, 2011.
- Garza Villarreal EA, Brattico E, Leino S, Ostergaard L, Vuust P.** Distinct neural responses to chord violations: a multiple source analysis study. *Brain Res* 1389: 103–114, 2011.
- Geschwind N, Levitsky W.** Human brain: left-right asymmetries in temporal speech region. *Science* 161: 186–187, 1968.
- Grahn JA, Brett M.** Rhythm and beat perception in motor areas of the brain. *J Cogn Neurosci* 19: 893–906, 2007.
- Griffiths TD, Warren JD.** The planum temporale as a computational hub. *Trends Neurosci* 25: 348–353, 2002.
- Griffiths TD, Johnsrude IS, Dean JL, Green GG.** A common neural substrate for the analysis of pitch and duration pattern in segmented sound? *Neuroreport* 10: 3825–3830, 1999.
- Griffiths TD, Rees A, Witton C, Cross PM, Shakir RA, Green GG.** Spatial and temporal auditory processing deficits following right hemisphere infarction. *Brain* 120: 785–794, 1997.
- Hoch L, Poulin-Charronnat B, Tillmann B.** The influence of task-irrelevant music on language processing: syntactic and semantic structures. *Front Psychol* 2: 112, 2011.
- Holmes AP, Friston KJ.** Generalisability, random effects and population inference. *Neuroimage* 7: S754, 1998.
- Humphries C, Love T, Swinney D, Hickok G.** Response of anterior temporal cortex to syntactic and prosodic manipulations during sentence processing. *Hum Brain Mapp* 26: 128–138, 2005.
- Jackendoff R, Lerdahl F.** The capacity for music: what is it, and what's special about it? *Cognition* 100: 33–72, 2006.
- Janata P.** ERP measures assay the degree of expectancy violation of harmonic contexts in music. *J Cogn Neurosci* 7: 153–164, 1995.
- Janata P, Grafton ST.** Swinging in the brain: shared neural substrates for behaviors related to sequencing and music. *Nat Neurosci* 6: 682–687, 2003.
- Jones MR, Boltz M.** Dynamic attending and responses to time. *Psychol Rev* 96: 459–491, 1989.
- Julian J, Fedorenko E, Webster J, Kanwisher N.** An algorithmic method for functionally defining regions of interest in the ventral visual pathway. *Neuroimage* 60: 2357–2364, 2012.
- Justus TC, Bharucha JJ.** Modularity in musical processing: the automaticity of harmonic priming. *J Exp Psychol Hum Percept Perform* 27: 1000–1011, 2001.
- Keller SS, Highley JR, Garcia-Finana M, Sluming V, Rezaie R, Roberts N.** Sulcal variability, stereological measurement and asymmetry of Broca's area in MR images. *J Anat* 211: 534–555, 2007.
- Koelsch S.** Neural substrates of processing syntax and semantics in music. *Curr Opin Neurobiol* 15: 207–212, 2005.
- Koelsch S.** Toward a neural basis of music perception—a review and updated model. *Front Psychol* 2: 110, 2011.
- Koelsch S, Siebel WA.** Towards a neural basis of music perception. *Trends Cogn Sci* 9: 578–584, 2005.
- Koelsch S, Fritz T, Schulze K, Alsop D, Schlaug G.** Adults and children processing music: an fMRI study. *Neuroimage* 25: 1068–1076, 2005.

- Koelsch S, Gunter TC, von Cramon DY, Zysset S, Lohmann G, Friederici AD. Bach speaks: a cortical "language-network" serves the processing of music. *Neuroimage* 17: 956–966, 2002.
- Koelsch S, Gunter TC, Schroger E, Tervaniemi M, Sammler D, Friederici AD. Differentiating ERAN and MMN: an ERP study. *Neuroreport* 12: 1385–1389, 2001.
- Kornysheva K, von Cramon DY, Jacobsen T, Schubotz RI. Tuning-in to the beat: aesthetic appreciation of musical rhythms correlates with a premotor activity boost. *Hum Brain Mapp* 31: 48–64, 2010.
- Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI. Circular analysis in systems neuroscience—the dangers of double dipping. *Nat Neurosci* 12: 535–540, 2009.
- Krizman J, Marian V, Shook A, Skoe E, Kraus N. Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. *Proc Natl Acad Sci USA* 109: 7877–7881, 2012.
- Krumhansl CL. Rhythm and pitch in music cognition. *Psychol Bull* 126: 159–179, 2000.
- Leino S, Brattico E, Tervaniemi M, Vuust P. Representation of harmony rules in the human brain: further evidence from event-related potentials. *Brain Res* 1142: 169–177, 2007.
- Lerdahl F, Jackendoff R. *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press, 1983.
- Levitin DJ, Menon V. Musical structure is processed in "language" areas of the brain: a possible role for Brodmann area 47 in temporal coherence. *Neuroimage* 20: 2142–2152, 2003.
- Levitin DJ, Menon V. The neural locus of temporal structure and expectancies in music: evidence from functional neuroimaging at 3 Tesla. *Music Percept* 22: 563–575, 2005.
- Liégeois-Chauvel C, Peretz I, Babai M, Laguitton V, Chauvel P. Contribution of different cortical areas in the temporal lobes to music processing. *Brain* 121: 1853–1867, 1998.
- Limb CJ. Structural and functional neural correlates of music perception. *Ann Rev* 288A: 435–446, 2006.
- Luria AR, Tsvetkova L, Futer DS. Aphasia in a composer. *J Neurol Sci* 1: 288–292, 1965.
- Maess B, Koelsch S, Gunter TC, Friederici AD. Musical syntax is processed in Broca's area: an MEG study. *Nat Neurosci* 4: 540–545, 2001.
- Malach R, Reppas JB, Benson R, Kwong KK, Jiang H, Kennedy WA, Ledden PJ, Brady TJ, Rosen BR, Tootell RB. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc Natl Acad Sci USA* 92: 8135–8139, 1995.
- McDermott J, Hauser MD. The origins of music: innateness, uniqueness, and evolution. *Music Percept* 23: 29–59, 2005.
- McDermott JH, Lehr AJ, Oxenham AJ. Individual differences reveal the basis of consonance. *Curr Biol* 20: 1035–1041, 2010.
- Merrill J, Sammler D, Bangert M, Goldhahn D, Lohmann G, Turner R, Friederici A. Perception of words and pitch patterns in song and speech. *Front Psychol* 3: 76, 2012.
- Miller EK, Cohen JD. An integrative theory of prefrontal cortex function. *Annu Rev Neurosci* 24: 167–202, 2001.
- Milner B. Laterality effects in audition. In: *Interhemispheric Relations and Cerebral Dominance*, edited by Mountcastle VB. Baltimore, MD: Johns Hopkins Press, 1962, p. 177–195.
- Moreno S, Bialystok E, Barac R, Schellenberg EG, Cepeda N, Chau T. Short-term music training enhances verbal intelligence and executive function. *Psychol Sci* 22: 1425–1433, 2011.
- Neville H, Andersson A, Bagdade O, Bell T, Currin J, Fanning J, Heidenreich L, Klein S, Lauinger B, Pakula E, Paulsen D, Sabourin L, Stevens C, Sundborg S, Yamada Y. How can musical training improve cognition? In: *The Origins of Human Dialog: Speech and Music*, edited by Dehaene S, Petit C. Paris: Odile Jacob, 2009, p. 277–290.
- Neville HJ, Bavelier D, Corina D, Rauschecker J, Karni A, Lalwani A, Braun A, Clark V, Jezzard P, Turner R. Cerebral organization for language in deaf and hearing subjects: biological constraints and effects of experience. *Proc Natl Acad Sci USA* 95: 922–929, 1998.
- Nieto-Castañón A, Fedorenko E. Functional localizers increase the sensitivity and functional resolution of multi-subject analyses. *Neuroimage* 22: 222–223, 2012.
- Nieto-Castañón A, Ghosh SS, Tourville JA, Guenther FH. Region of interest based analysis of functional imaging data. *Neuroimage* 19: 1303–1316, 2003.
- Norman-Haignere SV, McDermott JH, Kanwisher N. Regions of auditory cortex exhibit pitch-selective responses across a wide variety of sounds. *Advances and Perspectives in Auditory Neurophysiology*, Satellite Symposium at the Society for Neuroscience Annual Meeting, 2011.
- Ono M, Kubik S, Abernathy CD. *Atlas of the Cerebral Sulci*. Stuttgart, Germany: Thieme, 1990.
- Opitz B, Rinne T, Mecklinger A, von Cramon DY, Schroger E. Differential contribution of frontal and temporal cortices to auditory change detection: fMRI and ERP results. *Neuroimage* 15: 167–174, 2002.
- Osnes B, Hugdahl K, Hjelmervik H, Specht K. Stimulus expectancy modulates inferior frontal gyrus and premotor cortex activity in auditory perception. *Brain Lang* 22: 222–223, 2012.
- Parbery-Clark A, Tierney A, Strait DL, Kraus N. Musicians have fine-tuned neural distinction of speech syllables. *Neuroscience* 219: 111–119, 2012.
- Patel AD. Language, music, syntax and the brain. *Nat Neurosci* 6: 674–681, 2003.
- Patel AD. *Music, Language, and the Brain*. New York: Oxford Univ. Press, 2008.
- Patel AD, Gibson E, Ratner J, Besson M, Holcomb PJ. Processing syntactic relations in language and music: an event-related potential study. *J Cogn Neurosci* 10: 717–733, 1998a.
- Patel AD, Peretz I, Tramo M, Labrecque R. Processing prosodic and musical patterns: a neuropsychological investigation. *Brain Lang* 61: 123–144, 1998b.
- Patterson RD, Uppenkamp S, Johnsrude IS, Griffiths TD. The processing of temporal pitch and melody information in auditory cortex. *Neuron* 36: 767–776, 2002.
- Penagos H, Melcher JR, Oxenham AJ. A neural representation of pitch salience in nonprimary human auditory cortex revealed with functional magnetic resonance imaging. *J Neurosci* 24: 6810–6815, 2004.
- Peretz I. Auditory atonalia for melodies. *Cogn Neuropsychol* 10: 21–56, 1993.
- Peretz I, Coltheart M. Modularity of music processing. *Nat Neurosci* 6: 688–691, 2003.
- Peretz I, Hyde K. What is specific to music perception? Insights from congenital amusia. *Trends Cogn Sci* 7: 362–367, 2003.
- Peretz I, Zatorre RJ. Brain organization for music processing. *Annu Rev Psychol* 56: 89–114, 2005.
- Peretz I, Kolinsky R, Tramo M, Labrecque R, Hublet C, Demeurisse C, Belleville S. Functional dissociations following bilateral lesions of auditory cortex. *Brain* 117: 1283–1301, 1994.
- Pernet CR, Charest I, Belizaire G, Zatorre RJ, Belin P. The temporal voice areas: spatial characterization and variability. *Neuroimage* 36: S1–S168, 2007.
- Phillips-Silver J, Toivianen P, Gosselin N, Piche O, Nozaradan S, Palmer C, Peretz I. Born to dance but beat deaf: a new form of congenital amusia. *Neuropsychologia* 49: 961–969, 2011.
- Piccirilli M, Sciarra T, Luzzi S. Modularity of music: evidence from a case of pure amusia. *J Neurol Neurosurg Psychiatry* 69: 541–545, 2000.
- Rogalsky C, Rong F, Saberi K, Hickok G. Functional anatomy of language and music perception: temporal and structural factors investigated using fMRI. *J Neurosci* 31: 3843–3852, 2011.
- Sammler D, Koelsch S, Friederici AD. Are left fronto-temporal brain areas a prerequisite for normal music-syntactic processing? *Cortex* 47: 659–673, 2011.
- Samson F, Zeffiro TA, Toussaint A, Belin P. Stimulus complexity and categorical effects in human auditory cortex: an activation likelihood estimation meta-analysis. *Front Psychol* 1: 241, 2011.
- Schellenberg EG. Examining the association between music lessons and intelligence. *Br J Psychol* 102: 283–302, 2011.
- Slevc LR, Rosenberg JC, Patel AD. Making psycholinguistics musical: self-paced reading time evidence for shared processing of linguistic and musical syntax. *Psychon Bull Rev* 16: 374–381, 2009.
- Slevc LR. Language and music: sound, structure, and meaning. *Cogn Sci* 3: 483–492, 2012.
- Sluming V, Brooks J, Howard M, Downes JJ, Roberts N. Broca's area supports enhanced visuospatial cognition in orchestral musicians. *J Neurosci* 27: 3799–3806, 2007.
- Stalinski SM, Schellenberg G. Music cognition: a developmental perspective. *Top Cogn Sci* 4: 485–497, 2012.
- Steinke WR, Cuddy LL, Jakobson LS. Dissociations among functional subsystems governing melody recognition after right hemisphere damage. *Cogn Neuropsychol* 18: 411–437, 2001.
- Stevens CJ. Music perception and cognition: a review of recent cross-cultural research. *Top Cogn Sci* 4: 653–667, 2012.
- Stewart L, von Kriegstein K, Warren JD, Griffiths TD. Music and the brain: disorders of musical listening. *Brain* 129: 2533–2553, 2006.

- Strait DL, Kraus N.** Can you hear me now? Musical training shapes functional brain networks for selective auditory attention and hearing speech in noise. *Front Psychol* 2: 113, 2011.
- Strait DL, Hornickel J, Kraus N.** Subcortical processing of speech regularities underlies reading and music aptitude in children. *Behav Brain Funct* 7: 44, 2011.
- Tahmasebi AM, David MH, Wild CJ, Rodd JM, Hakyemez H, Abolmaesumi P, Johnsruide IS.** Is the link between anatomical structure and function equally strong at all cognitive levels of processing? *Cereb Cortex* 22: 1593–1603, 2012.
- Terhardt E.** The concept of musical consonance: a link between music and psychoacoustics. *Music Percept* 1: 276–295, 1984.
- Tervaniemi M, Szameitat AJ, Kruck S, Schroger E, Alter K, De Baene W, Friederici A.** From air oscillations to music and speech: functional magnetic resonance imaging evidence for fine-tuned neural networks in audition. *J Neurosci* 26: 8647–8652, 2006.
- Tillmann B.** Music and language perception: expectations, structural integration and cognitive sequencing. *Top Cogn Sci* 4: 568–584, 2012.
- Tillmann B, Koelsch S, Escoffier N, Bigand E, Lalitte P, Friederici AD, von Cramon DY.** Cognitive priming in sung and instrument music: activation of inferior frontal cortex. *Neuroimage* 31: 1771–1782, 2006.
- Tillmann B, Bharucha JJ, Bigand E.** Implicit learning of tonality: a self-organizing approach. *Psychol Rev* 107: 885–913, 2000.
- Tillmann B, Janata P, Bharucha JJ.** Activation of the inferior frontal cortex in musical priming. *Cogn Brain Res* 16: 145–161, 2003.
- Vandenberghe R, Nobre AC, Price CJ.** The response of left temporal cortex to sentences. *J Cogn Neurosci* 14: 550–560, 2002.
- Warrier CM, Zatorre RJ.** Right temporal cortex is critical for utilization of melodic contextual cues in a pitch constancy task. *Brain* 127: 1616–1625, 2004.
- Wilson SJ, Pressing JL, Wales RJ.** Modelling rhythmic function in a musician post-stroke. *Neuropsychologia* 40: 1494–1505, 2002.
- Wong PC, Skoe E, Russo NM, Dees T, Kraus N.** Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat Neurosci* 10: 420–422, 2007.
- Zatorre RJ, Belin P.** Spectral and temporal processing in human auditory cortex. *Cereb Cortex* 11: 946–953, 2001.
- Zatorre RJ, Schoenwiesner M.** Cortical speech and music processes revealed by functional neuroimaging. In: *The Auditory Cortex*, edited by Winer JA, Schreiner CE. New York: Springer, 2011.
- Zatorre RJ, Evans AC, Meyer E.** Neural mechanisms underlying melodic perception and memory for pitch. *J Neurosci* 14: 1908–1919, 1994.

