Inter-subject synchronization of brain responses during natural music listening

Daniel A. Abrams,[1] Srikanth Ryali,[1] Tianwen Chen,[1] Parag Chordia,[4] Amirah Khouzam,[1] Daniel J. Levitin[5] and Vinod Menon[1,2,3]

[1]Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, CA, 94304, USA
[2]Program in Neuroscience, Stanford University School of Medicine, Stanford, CA, USA
[3]Department of Neurology and Neurological Sciences, Stanford University School of Medicine, Stanford, CA, USA
[4]Department of Music, Georgia Institute of Technology, Atlanta, GA, USA
[5]Department of Psychology, McGill University, Montreal, QC, Canada

Abstract

Music is a cultural universal and a rich part of the human experience. However, little is known about common brain systems that support the processing and integration of extended, naturalistic 'real-world' music stimuli. We examined this question by presenting extended excerpts of symphonic music, and two pseudomusical stimuli in which the temporal and spectral structure of the Natural Music condition were disrupted, to non-musician participants undergoing functional brain imaging and analysing synchronized spatiotemporal activity patterns between listeners. We found that music synchronizes brain responses across listeners in bilateral auditory midbrain and thalamus, primary auditory and auditory association cortex, right-lateralized structures in frontal and parietal cortex, and motor planning regions of the brain. These effects were greater for natural music compared to the pseudo-musical control conditions. Remarkably, inter-subject synchronization in the inferior colliculus and medial geniculate nucleus was also greater for the natural music condition, indicating that synchronization at these early stages of auditory processing is not simply driven by spectro-temporal features of the stimulus. Increased synchronization during music listening was also evident in a right-hemisphere fronto-parietal attention network and bilateral cortical regions involved in motor planning. While these brain structures have previously been implicated in various aspects of musical processing, our results are the first to show that these regions track structural elements of a musical stimulus over extended time periods lasting minutes. Our results show that a hierarchical distributed network is synchronized between individuals during the processing of extended musical sequences, and provide new insight into the temporal integration of complex and biologically salient auditory sequences.

Introduction

Music is a cultural universal and a rich part of the human experience. Brain imaging studies have identified an array of structures that underlie critical components of music, including pitch (Zatorre *et al.*, 1994; Patel & Balaban, 2001), harmony (Janata *et al.*, 2002; Passynkova *et al.*, 2005), rhythm (Snyder & Large, 2005; Grahn & Rowe, 2009), timbre (Menon *et al.*, 2002; Deike *et al.*, 2004) and musical syntax (Levitin & Menon, 2005; Abrams *et al.*, 2011; Oechslin *et al.*, 2012). A drawback of probing neural substrates of individual musical features is that artificially constructed laboratory stimuli do not represent music as it is commonly heard, limiting the ecological validity of such studies. Furthermore, this componential approach fails to tap into one of the most important aspects of listeners' musicality – the ability to integrate components of musical information over extended time periods (on the order of minutes) into a coherent perceptual gestalt (Leaver *et al.*, 2009).

Examining the synchronization of brain responses across listeners constitutes a novel approach for exploring neural substrates of musical information processing. Inter-subject synchronization (ISS) using functional magnetic resonance imaging (fMRI) detects common stimulus-driven brain structures by calculating voxel-wise correlations in fMRI activity over time between subjects (Hasson *et al.*, 2004). The theoretical basis for using this approach is that brain structures that are consistently synchronized across subjects during an extended stimulus constitute core brain regions responsible for tracking structural elements of that stimulus over time (Hasson *et al.*, 2010).

ISS represents a fundamentally different approach, and provides advantages, relative to conventional fMRI methods (Wilson *et al.*, 2008; see Fig. S1). ISS allows us to examine cognitive processes that require the integration of information over extended time periods; this is critical for the study of music in which the structure of musical elements is manifested over time. Furthermore, ISS does not rely on *a priori* assumptions about specific stimulus events or subtraction paradigms that require comparison of discrete perceptual or cognitive events.

Our goal was to examine shared neural representations underlying the processing of natural musical stimuli ('Natural Music'; Fig. 1).

Correspondences: Dr Daniel A. Abrams and Dr Vinod Menon, [1]Department of Psychiatry & Behavioral Sciences, as above.
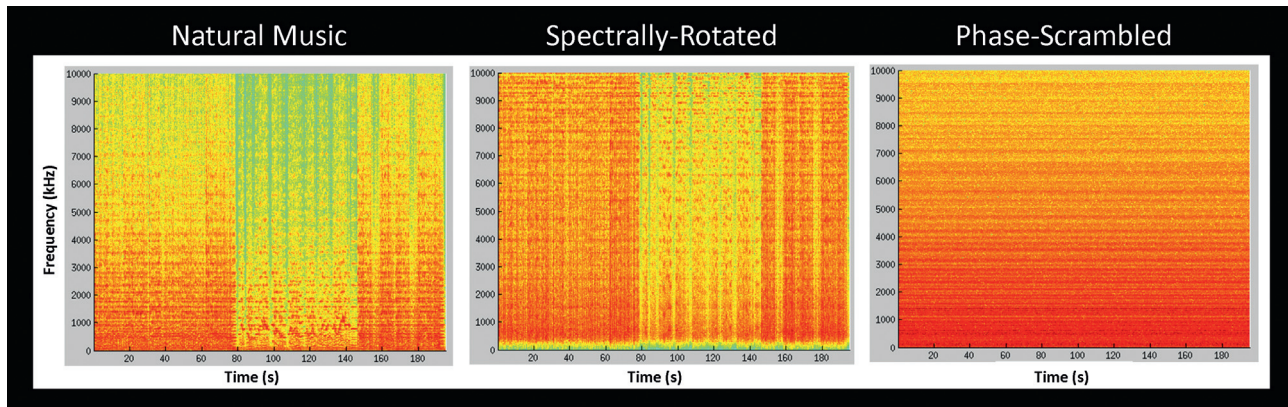E-mails: daa@stanford.edu and menon@stanford.edu

FIG. 1. Stimuli. Spectrograms of the Natural Music (left), Spectrally-Rotated (center) and Phase- Scrambled (right) conditions. The first of the four symphonies played during the fMRI scan is plotted for all conditions. Spectral-rotation and phase-scrambling was performed on each of the four symphonies comprising the stimulus set for the fMRI experiment.

We used ISS to identify brain regions that showed synchronized activity across individuals in response to music. To control for ISS that results from acoustical stimulus features as opposed to structural elements of the music stimulus, ISS results were compared with synchronization measured while subjects listened to two pseudo-musical stimuli in which the temporal and spectral structure of the Natural Music condition were disrupted ('Phase-Scrambled' and 'Spectrally-Rotated' stimuli; see Data S1). Consistent with previous findings (Joris *et al.*, 2004), we hypothesized that the presence of spectro-temporal modulations in the Spectrally-Rotated condition would drive consistent responses in auditory midbrain, thalamus and primary cortex while the absence of temporal modulations in the Phase-Scrambled condition would yield reduced ISS results in these structures. Importantly, we hypothesized that only the Natural Music condition would elicit ISS beyond primary sensory cortices into motor planning and fronto-parietal cortices, which underlie rhythmic (Chen *et al.*, 2008) and attentional processing (Sridharan *et al.*, 2007) of musical stimuli, respectively.

Materials and methods

Participants

The Stanford University School of Medicine Human Subjects committee approved the study, and informed consent was obtained from all participants. Seventeen right-handed subjects (nine males) between the ages of 19 and 27 years (mean = 21.3, SD = 1.78) with little or no musical training according to previously published criteria (Maess *et al.*, 2001) served as participants. The participants received $50 in compensation for participation.

Stimuli

Stimuli consisted of four symphonies of the late-baroque period composer William Boyce. Recordings were digitized at a sampling rate of 44.1 kHz in 16-bit mono. The total duration for these symphonies was 9 min 35 s. These particular symphonies were chosen for this study as they are representative of the Western music tradition yet they were unlikely to be recognized by the participants, thereby avoiding familiarity and memory-related effects. The four symphonies contained ten movement boundaries which were removed in order to ensure that event transitions were not driving ISS. To remove the movement boundaries, we first plotted each movement in Matlab and visually

identified when the final note of the movement descended into the noise floor of the recording. All subsequent samples beyond this point were removed from the movement. We evaluated each movement boundary removal by listening to the manipulated stimuli and ensuring that the final note of each movement was completely audible and decayed naturally. All silent samples at the beginning of each movement were removed using the same visual and auditory-guided procedures. The result of this manipulation was a seamless transition from movement to movement that lacked the relatively long periods of silence (~5 s) that characterize natural movement boundaries. The task was programmed with E-Prime (PSTNET, Pittsburgh, PA, USA; www.pstnet.com), and stimuli were presented binaurally at a comfortable listening level with noise-reducing headphones and a custom-built magnet-compatible audio system.

We used a freely available algorithm to perform spectral rotation on the musical stimuli (http://www.fil.ion.ucl.ac.uk/~jcrinion/rotation/blesser3.m). This method has been described in previous works (Blesser, 1972; Scott *et al.*, 2000; Warren *et al.*, 2006; Abrams *et al.*, 2012). The center frequency for spectral rotation was 5512 Hz. This center frequency was chosen so that the rotated frequencies would be within the frequency response range of the fMRI-compatible headphones (20–10 000 Hz). Phase-scrambling was performed by applying a Fourier transform to each of the four symphonies that constitute the Natural Music stimulus and then randomizing its phase response by adding a random phase shift at every frequency (Prichard & Theiler, 1994). The phase shifts were obtained by randomly sampling in the interval $(0, 2\pi)$. This process preserves the power spectrum of each of the four symphonies. Note that, by design, the Phase-Scrambled control stimulus preserves spectral density but not time-dependent fluctuations. We preferred this design as it facilitates a simple and interpretable result: brain structures that show greater ISS for Natural Music compared with the Phase-Scrambled condition are sensitive to the temporal structure of music. Our design therefore forms a necessary starting point for future investigations of more complex time-dependent attributes of musical structure that lead to synchronized responses among subjects, perhaps using a wavelet transform that preserves both the spectral density and the time-dependent fluctuations in that density.

fMRI data acquisition

Brain images were acquired on a 3T GE Signa scanner using a standard GE whole head coil (software Lx 8.3). For the Natural Music,

Spectrally-Rotated and Phase-Scrambled conditions, images were acquired every 2 s in two runs that lasted 9 min 42 s. The sequence of these stimulus conditions was consistent across listeners: the Natural Music condition was presented first, the Phase-Scrambled condition was presented second and the Spectrally-Rotated condition was presented third. While it would have been preferable to have randomized the stimulus presentation order across subjects to control for attention and fatigue, we do not believe that this had a significant effect on the results given that there was vastly greater ISS for the final stimulus condition (Spectral-Rotation) relative to the penultimate stimulus condition (Phase-Scrambled), which would not have occurred had fatigue and attention negatively affected ISS results. Subjects were instructed to attend to all the music and music-like stimuli. To allow for a natural listening experience, we did not provide any additional instructions to the subjects. A custom-built head holder was used to prevent head movement. Twenty-eight axial slices (4.0 mm thick, 0.5 mm skip) parallel to the AC–PC line and covering the whole brain were imaged using a T2*-weighted gradient echo spiral pulse sequence (TR = 2000 ms, TE = 30 ms, flip angle = 70°), providing an in-plane spatial resolution of 3.125 mm (Glover & Lai, 1998). Images were reconstructed by gridding interpolation and inverse Fourier transform for each time point into $64 \times 64 \times 28$ image matrices (voxel size $3.125 \times 3.125 \times 4.5$ mm). fMRI data acquisition was synchronized to stimulus presentation using a TTL pulse sent by E-Prime to the scanner timing board.

fMRI data analysis

Preprocessing

fMRI data were preprocessed using SPM8 (www.fil.ion.ucl.ac.uk/spm/software/spm8). Images were realigned to correct for motion, corrected for errors in slice-timing, spatially transformed to standard stereotaxic space [based on the Montreal Neurologic Institute (MNI) coordinate system], resampled every 2 mm using sinc interpolation and smoothed with a 6-mm full-width half-maximum Gaussian kernel to decrease spatial noise prior to statistical analysis. Translational movement in millimeters (x, y, z) and rotational motion in degrees (pitch, roll, yaw) was calculated based on the SPM8 parameters for motion correction of the functional images in each participant. Confounding effects of fluctuations in global mean were removed by calculating the mean signal across all voxels for each time point and regressing out these values at the corresponding time points at each voxel in the brain. Controlling for the global mean is commonly performed in inter-subject correlation studies (Hasson *et al.*, 2004; Wilson *et al.*, 2008). To remove pre-processing artifacts and nonlinear saturation effects, we excluded the first six time points of the experiment from the analysis.

Inter-subject synchronization

The inter-subject correlation analysis was performed using the WFU BPM toolbox (www.fmri.wfubmc.edu/cms/software). Synchronization was calculated by computing Pearson correlations between the voxel time series in each pair of subjects (136 subject-to-subject comparisons total; see Fig. S2). Pearson correlation coefficients at each voxel were converted into Z-scores using Fisher transformation. We computed the Z-normalized group correlation map for each stimulus condition by performing a one-sample t-test at each voxel, using the Z-scores from each subject-to-subject comparison.

Differences between the general linear model (GLM) and ISS

The GLM identifies brain regions that have consistently greater univariate activity for music relative to rest measured across subjects. A significant limitation of GLM analysis is that it cannot identify brain structures that show highly consistent patterns of fMRI activity measured across subjects (Hasson *et al.*, 2010). Nevertheless, the great consistency of these patterns of activity across subjects, facilitated by ISS analysis, strongly suggests that these brain regions track aspects of musical structure across time that represent functionally important regions for the processing of naturalistic musical stimuli. Due to the continuous nature of the musical stimuli in the current study, a GLM analysis, which relies on comparison of fMRI activity across short-duration task conditions, was not possible.

Non-parametric thresholding

As the resulting 136 Z-transformed correlation coefficients are not independent, we used a permutation test (Nichols & Holmes, 2002) to derive a spatial extent threshold to correct for multiple comparisons. To comply with the construction of the original paired t-test, we formed two paired groups for each permutation. For one reconstructed group, we correlated one subject from the Natural Music condition, denoted as $Sub_{i,1}$, with a different subject from the Phase-Scrambled condition, denoted as $Sub_{j,2}$, where i and j represent subjects, 1 represents the Natural Music condition, and 2 represents the Phase-Scrambled condition. Correspondingly, for the paired Z-transformed correlation coefficient in the other reconstructed group, we correlated $Sub_{i,2}$ with $Sub_{j,1}$ (i.e. the same paired subjects but with switched conditions). We randomly paired subjects from different conditions 136 times to resemble the original 136 correlations between 17 subjects within the same condition. Similarly, a t statistic was constructed using a paired group t-test with 136 Z-transformed correlation coefficients. We repeated the same permutation procedure 80 times and derived an appropriate spatial extent threshold based on the maximum cluster size to control family-wise error under 5% with a voxel-wise P value < 0.005 based on a t-distribution with a degree of freedom of 135. The resulting spatial extent threshold was determined to be 50 voxels.

These particular values were used to threshold the Z-normalized group correlation map. To compare ISS results between stimulus conditions, we used the Z-scores at each voxel generated during the ISS analysis (see above) to calculate a difference map. Specifically, we subtracted Z-scores for the Spectrally-Rotated and Phase-Scrambled conditions from Z-scores from the Natural Music condition for each subject-to-subject comparison (136 subject-to-subject comparisons in total). This analysis was restricted to the voxels which showed suprathreshold ISS in the group correlation map for the Natural Music condition. Group t-maps for the (Natural Music minus Spectrally-Rotated) and (Natural Music minus Phase-Scrambled) comparisons were then computed by performing one-way t-tests across all 136 difference maps for each comparison. Group difference t-maps were then thresholded using the permutation test as described previously (P < 0.005 height; P < 0.05, 50 voxels extent). While our analysis and interpretation focuses on comparison of ISS differences between the Natural Music and the two control conditions, for the sake of completeness we have also presented synchronization maps associated with the Natural Music, Spectrally-Rotated and Phase-Scrambled conditions.

ISS in subcortical structures

To examine whether sub-cortical auditory structures, including the inferior colliculus (IC) of the midbrain and medial geniculate nucleus (MGN) of thalamus, showed differences in ISS for the Natural Music condition compared with the Spectrally-Rotated and Phase-Scrambled conditions, we used the Z-scores generated during the ISS analysis (see above) to calculate the difference in Z-scores between the Natural Music and the control conditions within these regions of interest (ROIs). Specifically, we subtracted Z-scores for the Spectrally-Rotated and Phase-Scrambled conditions from Z-scores from the Natural Music condition for each subject-to-subject comparison (136 subject-to-subject comparisons in total). This analysis was restricted to the voxels within the IC and MGN as reported in a previous MRI study (Muhlau *et al.*, 2006). Based on the coordinates reported in that study, we used a sphere with a radius of 5 mm centered at ± 6, –33, –11 for the inferior colliculus ROIs and a sphere with a radius of 8 mm centered at ± 17, –24, –2 for the medial geniculate ROI. Given the relatively small sizes of these subcortical structures (5- and 8-mm spheres for the IC and MGN, respectively), the resulting difference Z-scores were thresholded at $P < 0.05$, uncorrected for extent.

Consistency and potential confounds in ISS

We performed three additional analyses to examine the possibility that our ISS results did not arise from stimulus-following, spectro-temporally invariant neural responses and synchronized inter-subject movement. First, we performed a within-subject analysis to examine whether neural activity measured across ROIs identified with ISS represents a global, uniform signal as opposed to regionally specific processing. We reasoned that if ISS represents either stimulus-following or consistent responses at each time point, fMRI time courses would be similar across all ROIs. To isolate neural activity from specific brain regions, we first created ROIs by crossing the thresholded ISS map for the Natural Music condition with eight right-hemisphere auditory and non-auditory cortical ROIs from the Harvard–Oxford probabilistic structural atlas, including Heschl's gyrus (HG), planum temporale (PT), planum polare (PP), posterior superior temporal gyrus (pSTG), BA 45 (extending into BA 47), posterior supramarginal gyrus (pSMG), mid-cingulate cortex (MCC) and pre-central gyrus (Smith *et al.*, 2004). A probability threshold of 25% was used to define each anatomical ROI in the Harvard–Oxford probabilistic structural atlas, and these thresholded ROIs were binarized prior to additional processing. We also included the two sub-cortical auditory ROIs described previously as well as the PGa and PGp sub-divisions of the angular gyrus (AG; Caspers *et al.*, 2006), resulting in a total of 12 ROIs. We then extracted the time-series for each ROI and subject for all three stimulus conditions, measured as the first principal eigenvector from all voxels within each ROI. The 12 ROI-specific time series were then correlated on a within-subject basis, resulting in 66 region-to-region Pearson correlation values for each subject. The resulting Pearson's correlation values were converted to Z-scores using the Fisher transform. To perform group statistics, we calculated one-sample *t*-tests on the Fisher-transformed correlation values for each region-to-region connection measured across subjects. The *t*-test results were FDR corrected using a threshold of $P < 0.01$.

In the second analysis, the goal was to examine whether significant ISS during the Natural Music condition was associated with constant synchronization of subjects' fMRI time-series measured across the entire musical sequence, or alternatively whether ISS was associated with isolated and concentrated periods of synchronization measured in the musical sequence. To this end, we performed an inter-subject time-frequency analysis using a continuous wavelet transform in order to examine changes in synchronization over time and frequency (Torrence & Compo, 1998; Grinsted *et al.*, 2004). In this analysis, we computed the wavelet cross spectra between ROI time series extracted from all pairs of subjects at 64 different frequency scales using the Matlab function 'wcoher.m' (www.mathworks.com/products/matlab) with 'cgau2' as a mother wavelet. The wavelet cross spectrum C_{xy} of two time series x and y is defined as:

$$C_{xy} = S(C_x^*(a,b)C_y(a,b))$$

where $C_x(a,b)$ and $C_y(a,b)$ denote continuous wavelet transforms of x and y at scales a and positions b. The superscript * is the complex conjugate and S is a smoothing operator in time and scale. The time series entered into this analysis were the same time series used in the previous analysis, which is the first eigenvector calculated across all voxels within each ROI.

In the third analysis, the goal was to examine whether correlations in subjects' movement patterns within the scanner may have driven ISS results. To address this question, we performed an inter-subject correlation analysis using the time series for each of the six movement parameters. Similar to the main ISS analysis described previously, we calculated Pearson's correlations for all pair-wise subject comparisons (i.e. 136 subject-to-subject comparisons) for each of the six time-varying movement parameters specified by SPM8 during fMRI data pre-processing (i.e. *x*, *y*, *z*, pitch, roll, yaw) for both the Natural Music and the Phase-Scrambled conditions. Data were linearly detrended prior to performing the correlation analysis. The resulting Pearson's correlation values for all subject-to-subject comparisons were Fisher transformed, and then these values were entered into a paired *t*-test (i.e. Natural Music vs. Phase-Scrambled) to examine whether movement correlations measured during the Natural Music condition were significantly different from those measured during the Phase-Scrambled condition.

Results

Inter-subject synchronization

We measured fMRI activity in 17 adult non-musicians while they listened to 9.5 min of symphonic music from the late-Baroque period and the Spectrally-Rotated and Phase-Scrambled versions of those same compositions (control stimuli). Musical stimuli were similar to those used in a previous study investigating neural dynamics of event segmentation in music across the boundaries of musical movements (Sridharan *et al.*, 2007), except that here we removed 'silent' movement boundaries from the musical stimuli. This stimulus manipulation enabled us to isolate brain synchronization during audible musical segments. We found that a highly distinctive and distributed set of brain regions was synchronized between subjects during Natural Music listening (Table 1), including subcortical and cortical auditory structures as well as structures in frontal, parietal and insular cortices.

ISS in sub-cortical auditory structures

Examination of ISS maps for the Natural Music condition showed that synchronization was evident throughout the right-hemisphere IC

TABLE 1. Peak ISS Z-Scores for the Natural Music condition

Brain region	Maximum Z-score, Natural Music	x	y	z
Left IC	3.038	−2	−34	−10
Right IC	4.920	6	−32	−8
Right MGN	3.721	14	−28	4
Left HG	12.831	−48	−12	0
Left PP	11.269	−48	−14	−2
Left PT	10.538	−42	−32	8
Left pSTG	5.740	−66	−20	10
Right HG	8.441	52	−14	2
Right PP	8.868	46	−16	0
Right PT	8.436	60	−18	4
Right pSTG	9.566	62	−18	2
Left PGa	5.834	−28	−72	48
Left PGp	5.756	−54	−64	22
Right BA 44	7.183	50	18	2
Right BA 45	7.013	50	20	2
Right BA 47	7.494	46	20	−6
Right SMG	3.977	58	−40	6
Right PGa	5.305	54	−48	22
Right PGp	2.957	42	−62	30
Left MCC	8.668	−2	−28	46
Right MCC	7.770	10	−26	44
Right Precentral/PMC	7.475	12	−26	44

FIG. 2. Inter-subject synchronization in subcortical auditory structures. (A) Axial slices (Z = −11) reveal ISS in the inferior colliculus (IC) of the midbrain in response to the Natural Music (left) but not to the Spectrally-Rotated (center) and Phase-Scrambled music (right) conditions. Images were thresholded using a voxel-wise statistical height threshold of ($P < 0.005$), with corrections for multiple comparisons at the cluster level ($P < 0.05$; 50 voxels). (B) Results show suprathreshold voxels throughout the IC (top) and MGN (bottom) for the Natural Music > Spectrally-Rotated (left) and Natural Music > Phase-Scrambled (right) comparisons. Sub-cortical ROIs were thresholded using a voxel-wise statistical height threshold of ($P < 0.05$), uncorrected. Functional images are superimposed on a standard brain from a single normal subject (MNI_152_T1_1mm_brain.nii; MRIcroN (Rorden & Brett, 2000)).

of the midbrain with a small extent evident in the left-hemisphere IC (Fig. 2A, left). Surprisingly, very little synchronization was evident in the IC for the Spectrally-Rotated and Phase-Scrambled control conditions (Fig. 2A, center and right). Furthermore, in a direct comparison of synchronization between the music and control conditions, we found significantly greater ISS for the Natural Music condition than for the control conditions throughout bilateral IC (Fig. 2B, top row). Based on this finding, we examined whether this effect was also evident in the MGN of the thalamus. Again, we found significantly greater ISS in the MGN for Natural Music relative to the control conditions (Fig. 2B, bottom row).

ISS in superior temporal cortex

The Natural Music condition also showed widespread synchronization in auditory cortex (Fig. 3, left), extending bilaterally from HG, which contains primary auditory cortex, into PP, PT and pSTG in auditory association cortex. Results for the Spectrally-Rotated condition also indicated widespread ISS in auditory cortical regions similar to the Natural Music condition (Fig. 3, center), although ISS results for the Phase-Scrambled condition showed that no auditory cortical voxels had significant synchronization (Fig. 3, right). This pattern was also evident when we directly compared synchronization between stimulus conditions. Specifically, there was no difference between auditory cortical synchronization for Natural Music and the Spectrally-Rotated conditions (Fig. 4, left) while there was significantly greater ISS for Natural Music compared with the Phase-Scrambled condition throughout each of these auditory cortical regions except for right-hemisphere HG and left-hemisphere pSTG (Fig. 4, right). This finding strongly suggests that temporal patterns present in Natural Music are necessary to drive ISS in auditory cortical regions.

ISS in fronto-parietal cortex

Synchronization for the Natural Music condition extended beyond auditory regions and into a variety of cortical regions associated
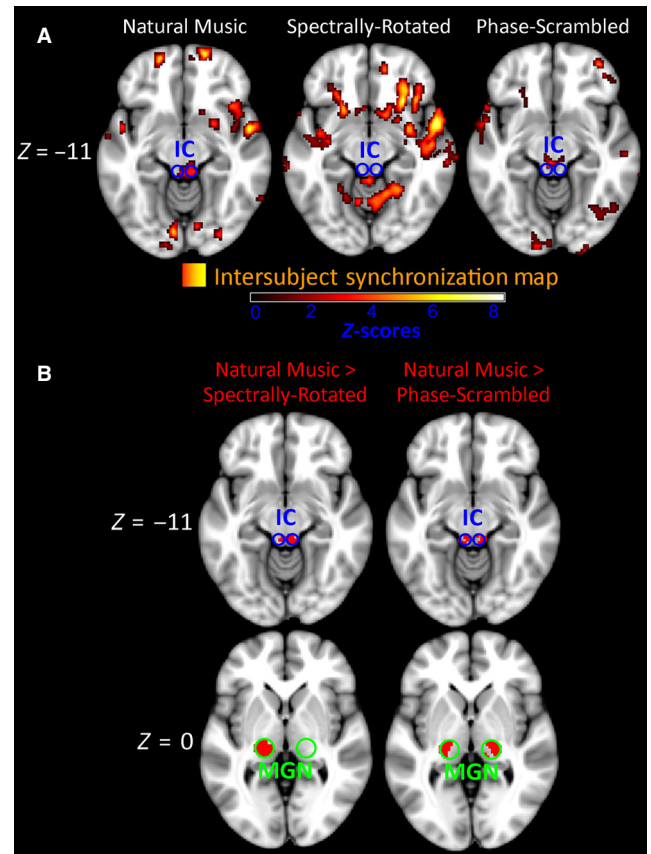
with higher-level cognitive function. First, ISS for Natural Music was evident in the right-hemisphere inferior frontal gyrus (IFG), including BA 45 and 47 (Fig. 5, top left). There was no suprathreshold ISS in the left hemisphere in either of these frontal structures. Additionally, ISS for the Natural Music condition was evident in multiple regions of the parietal lobe, including the PGa subdivision of the AG bilaterally, with a strong right-hemisphere bias, as well as the intra-parietal sulcus (IPS; Fig. 5, bottom left). In contrast to the Natural Music condition, the Spectrally-Rotated and Phase-Scrambled conditions resulted in significantly reduced synchronization across these fronto-parietal brain regions. For example, ISS for the Spectrally-Rotated condition showed only small extents in both the IFG (Fig. 5, top center) and PGa subdivision of the AG in the parietal cortex (Fig. 5, bottom center) and the Phase-Scrambled condition failed to induce ISS in either the IFG or the PGa (Fig. 5, right top and bottom). Direct comparisons between Natural Music and two control conditions indicated significantly greater synchronization in right-hemisphere BA 45 and 47 as well as PGa and IPS (Fig. 6), regions that we previously found to be involved in tracking temporal structure (Levitin & Menon, 2003).
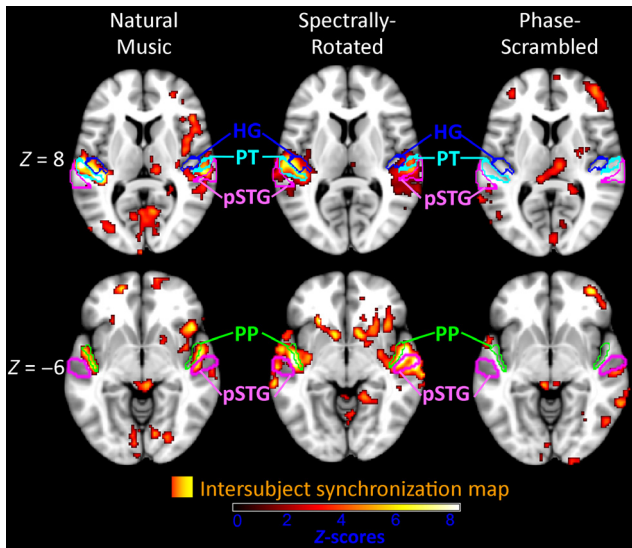
FIG. 3. Inter-subject synchronization in auditory cortex. Axial slices showing ISS for the Natural Music (left), Spectrally-Rotated (center) and Phase-Scrambled (right) conditions in dorsal ($Z = 8$; top) and ventral ($Z = -6$; bottom) views of auditory cortex. Results indicate ISS throughout auditory cortex, including Heschl's gyrus (HG, blue), planum temporale (PT, cyan), posterior superior temporal gyrus (pSTG, pink), and planum polare (PP, green), for the Natural Music and Spectrally-Rotated conditions but not for the Phase-Scrambled condition. Images were thresholded using a voxel-wise statistical height threshold of ($P < 0.005$), with corrections for multiple comparisons at the cluster level ($P < 0.05$; 50 voxels).

FIG. 5. Inter-subject synchronization in fronto-parietal cortex. Coronal slices showing ISS for the Natural Music (left), Spectrally-Rotated (center) and Phase-Scrambled (right) conditions in anterior ($Y = 20$; top) and posterior ($Y = -50$; bottom) views of the brain. Results indicate ISS for Natural Music in right hemisphere IFG, including BA 45 and 47, and parietal cortex, including the PGa subregion of the angular gyrus and the superior parietal lobule (SPL). ISS was greatly reduced across these frontal and parietal regions for both the Spectrally-Rotated and Phase-Scrambled control conditions. Images were thresholded using a voxel-wise statistical height threshold of ($P < 0.005$), with corrections for multiple comparisons at the cluster level ($P < 0.05$; 50 voxels).
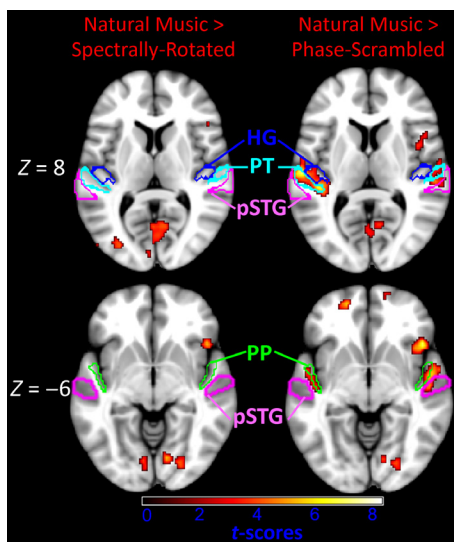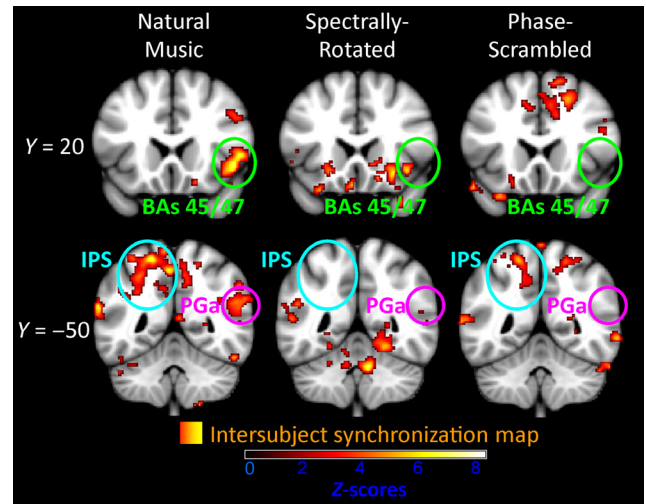

FIG. 4. ISS Difference maps in auditory cortex. Images show ISS difference maps for Natural Music > Spectrally-Rotated (left) and Natural Music > Phase-scrambled (right) comparisons. Results show no significant differences across auditory cortex for the Natural Music > Spectrally-Rotated comparison, but many suprathreshold voxels across these regions for the Natural Music > Phase-Scrambled comparison. Difference maps were thresholded using a voxel-wise statistical height threshold of ($P < 0.005$), with corrections for multiple comparisons at the cluster level ($P < 0.05$; 50 voxels).
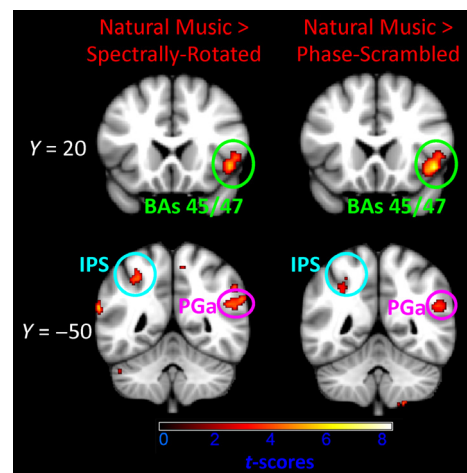

FIG. 6. ISS difference maps in fronto-parietal cortex. Images show ISS difference maps in frontal and parietal cortex for Natural Music > Spectrally-Rotated (left) and Natural Music > Phase-Scrambled (right) comparisons. Results show significant differences in BAs 45 and 47 of IFG (top), as well as PGa and IPS of parietal cortex (bottom), for both stimulus comparisons. Images were thresholded using a voxel-wise statistical height threshold of ($P < 0.005$), with corrections for multiple comparisons at the cluster level ($P < 0.05$; 50 voxels).

ISS in motor cortex

The Natural Music condition also revealed significant ISS in motor systems of the brain. Specifically, a functional cluster was identified in the premotor motor cortex (PMC), MCC and supplementary motor area, key cortical areas for movement planning, as well as the motor cortex bilaterally for the Natural Music condition (Fig. 7A, left). ISS for the Natural Music condition was also evident in the cerebellum in bilateral lobes VI and VIIb. ISS in response to the control conditions revealed smaller extents in these frontal motor regions (Fig. 7A, center and right), and the Phase-Scrambled condition failed to reveal ISS in any subregion of the

cerebellum. Direct comparison between the Natural Music and the control conditions revealed significantly greater ISS in the PMC in the right hemisphere and the MCC in both hemispheres (Fig. 7B). Moreover, there was greater ISS for Natural Music compared than for the Phase-Scrambled condition in left hemisphere lobe VI of the cerebellum.

Consistency of fMRI responses and potential confounds in ISS

A final goal of this work was to examine consistency of fMRI activity over time and, in doing so, investigate potential confounds that could influence our interpretation of ISS. Specifically, we examined several factors that would introduce high levels of ISS due to influences unrelated to music information processing. We reasoned that ISS confounds could arise from: (1) a 'low-level' stimulus-following response to the extended musical sequence rather than regionally specific brain processing of the musical stimulus, resulting in highly correlated fMRI activity patterns measured across auditory, motor and fronto-parietal brain regions; (2) invariant inter-subject correlation magnitudes measured over time during the extended Natural Music sequence, reflecting a consistent and static neural process dri-

FIG. 7. Inter-subject synchronization in motor-planning cortical regions. (A) Coronal (Y = −7) and sagittal (X = 13) slices shows ISS throughout the pre-motor cortex (PMC) and mid-cingulate cortex (MCC), respectively, in response to the Natural Music (left) condition. ISS was less prevalent in the both of these motor-planning regions for the Spectrally-Rotated (center) and Phase-Scrambled (right) condition. (B) Results show suprathreshold voxels in the right PMC and MCC for the Natural Music > Spectrally-Rotated and Natural Music > Phase-Scrambled comparisons. Images were thresholded using a voxel-wise statistical height threshold of (P < 0.005), with corrections for multiple comparisons at the cluster level (P < 0.05; 50 voxels).

ven by temporal regularities in the stimulus; or (3) synchronized subject movement during fMRI scanning that results in artifactual increases in the correlation of fMRI time-series measured for the Natural Music condition.

We performed three separate analyses to address these issues. First, to examine homogeneity of responses measured across the brain, we extracted fMRI time series for the Natural Music condition from 12 ROIs highlighted in the ISS results and performed a within-subject correlation analysis (see Methods). We hypothesized that stimulus-following would result in significant correlations in many (or most) of the 66 region-to-region comparisons. We found that less than that 20% of the inter-regional comparisons were significantly correlated, indicating that most regional neural activity in response to the Natural Music condition is highly specific and is not represented by a uniform, undifferentiated neural signal (Table 2). Importantly, results from the inter-regional analysis highlight the hierarchical structure of the auditory system during the processing of Natural Music. For example, significant positive connectivity in subcortical structures was specific to well-described connections in the ascending auditory system, including the IC to MGN connection as well as the MGN to HG connection (Kaas & Hackett, 2000). Additionally, the results also indicated highly synchronized responses among auditory cortical regions of superior temporal cortex, including HG, PP, PT and pSTG. The inter-regional analysis also identified three positively correlated long-range connections, including HG to IFG, HG to SMG, the fronto-parietal IFG to SMG connection, as well as one negatively correlated long-range connection, PP to the PGp division of the AG. We also examined inter-regional synchronization for the two control conditions using the same ROIs used for the Natural Music condition (Tables 3 and 4). The results show that inter-regional synchronization is similar between the Natural Music and Spectrally-Rotated conditions but, consistent with ISS results, inter-regional synchronization is sharply reduced in the Phase-Scrambled control condition. These results also provide novel evidence that ISS is distinct from inter-regional synchronization and represents fundamentally different aspects of information processing.

In the second analysis, we performed an inter-subject, cross-spectra analysis using a continuous wavelet transform to examine time-dependent, frequency-specific correlations between subjects' fMRI activity measured throughout the entire Natural Music stimulus (> 9 min in duration). We hypothesized that if the rhythm of the Natural Music, or any other temporal regularities evident in all subjects' fMRI data, was driving ISS results, then the cross-spectra magnitude would show consistently high amplitudes over time in subject-to-subject comparisons. The cross-spectra analysis revealed that correlations between subjects' fMRI time series from three right hemisphere ROIs (IC, HG and IFG) failed to show consistently high amplitudes over time (Fig. 8). Rather, intermittent and isolated periods of spectral coherence over time were evident, suggesting that consistent temporal regularities in the stimulus were not responsible for driving our observed ISS results.

In the third analysis, we examined whether consistent patterns of movement in the scanner may have driven ISS results. Here, we compared ISS (136 subject-to-subject comparisons) for the Natural Music and Phase-Scrambled conditions using the time series from the six affine movement parameters. Movement parameters did not differ (P > 0.3 for all movement parameters) between the Natural Music and Phase-Scrambled conditions, suggesting that consistent movement patterns across subjects induced by musical rhythm did not drive increased ISS in the Natural Music condition.
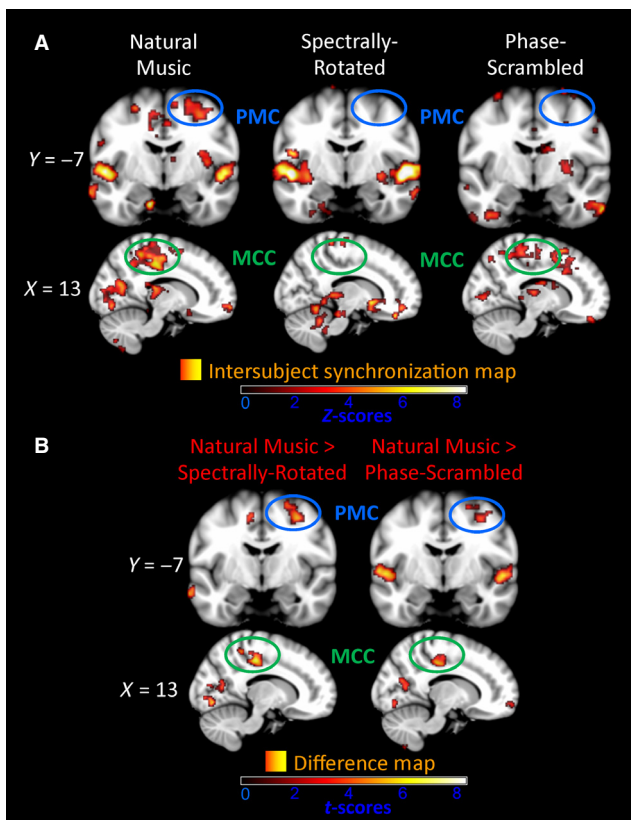
TABLE 2. Inter-regional synchronization for the Natural Music condition

	Right IC	Right MGN	Right HG	Right PP	Right PT	Right pSTG	Right IFG	Right PGa	Right PGp	Right SMG	Right PCC	Right Precentral
Right IC												
Right MGN	$R = 0.250$											
Right HG		$R = 0.251$										
Right PP			$R = 0.545$									
Right PT			$R = 0.503$	$R = 0.411$								
Right pSTG			$R = 0.293$	$R = 0.403$	$R = 0.475$							
Right IFG			$R = 0.277$									
Right PGa												
Right PGp				$R = -0.226$								
Right SMG			$R = 0.214$				$R = 0.248$					
Right PCC												
Right Precentral												

TABLE 3. Inter-regional synchronization for the Spectrally-Rotated condition

	Right IC	Right MGN	Right HG	Right PP	Right PT	Right pSTG	Right IFG	Right PGa	Right PGp	Right SMG	Right PCC	Right Precentral
Right IC												
Right MGN	$R = 0.178$											
Right HG												
Right PP			$R = 0.567$									
Right PT			$R = 0.708$	$R = 0.427$								
Right pSTG			$R = 0.348$	$R = 0.406$								
Right IFG			$R = 0.333$	$R = 0.409$	0.26025							
Right PGa												
Right PGp				$R = -0.156$				$R = 0.306$				
Right SMG							$R = 0.215$					
Right PCC												
Right Precentral				$R = 0.179$								

TABLE 4. Inter-regional synchronization for the Phase-Scrambled condition

	Right IC	Right MGN	Right HG	Right PP	Right PT	Right pSTG	Right IFG	Right PGa	Right PGp	Right SMG	Right PCC	Right Precentral
Right IC												
Right MGN												
Right HG												
Right PP			$R = 0.634$									
Right PT			$R = 0.545$									
Right pSTG			$R = 0.430$									
Right IFG												
Right PGa												
Right PGp												
Right SMG								$R = 0.386$				
Right PCC												
Right Precentral					$R = 0.262$							

Discussion

A complete understanding of human brain function requires the use of biologically realistic stimuli (Hasson *et al.*, 2010). We applied this principle to the study of music processing in the brain and identified a distributed network of brain regions that is synchronized across participants during Natural Music listening. This network includes sub-cortical and cortical auditory structures of the temporal lobe, inferior prefrontal cortex and parietal regions associated with attention and working memory, and medial frontal regions associated with motor planning. Nearly all of these brain structures have been implicated in some aspect of music processing in previous research (Zatorre *et al.*, 1994; Maess *et al.*, 2001; Janata *et al.*, 2002; Menon *et al.*, 2002; Snyder & Large, 2005), but the current results implicate these regions in the shared tracking of structural elements of music over extended time periods. Control conditions consisted of a Spectrally-Rotated condition, which contained the temporal features of the Natural Music condition but whose spectral features were rearranged relative to Natural Music, and a Phase-
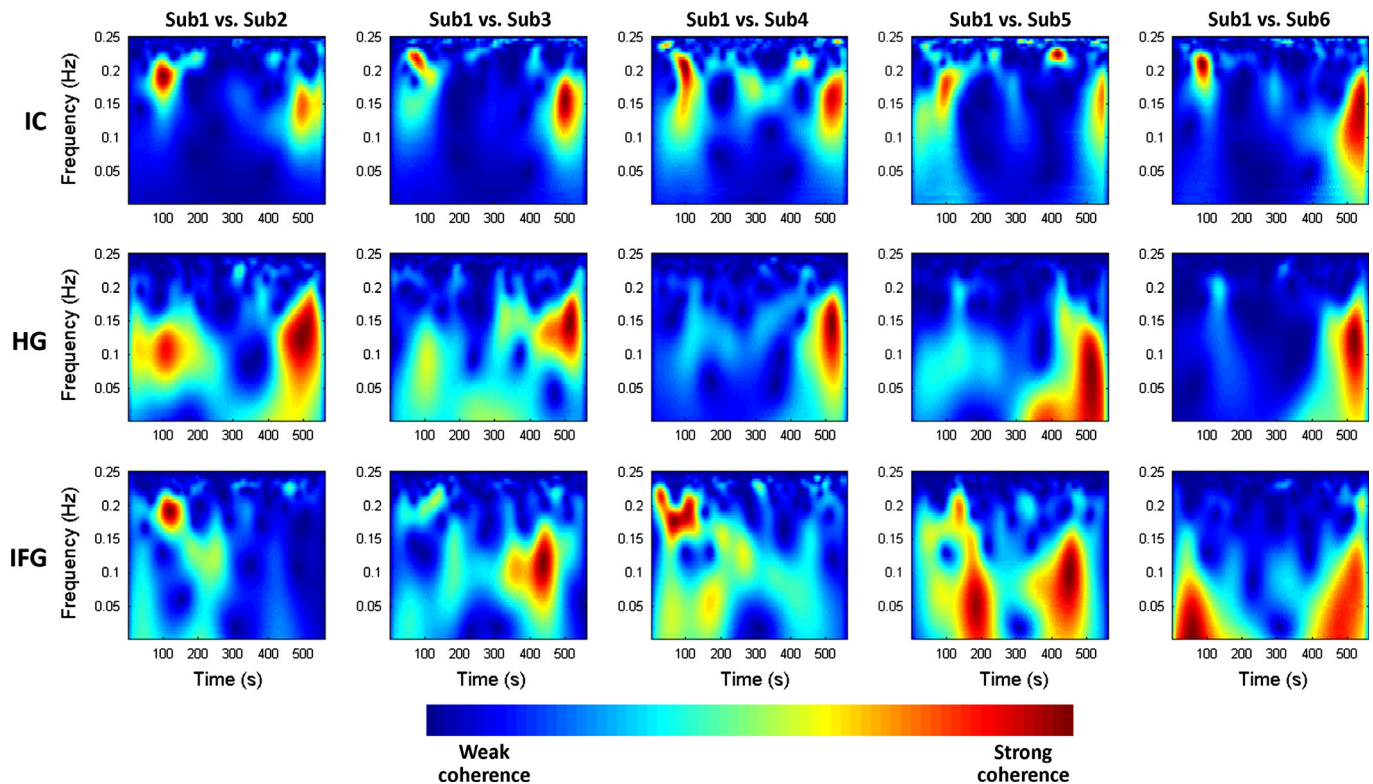
FIG. 8. Inter-subject spectral coherence analysis in three brain regions. Representative examples of five pairs of subject-to-subject cross-spectra during the Natural Music condition in the IC (top row), HG (middle row), and IFG (bottom row). Intermittent and isolated periods of spectral coherence over time were observed, indicating that ISS does not arise from spectro-temporally invariant neural responses and stimulus-following. Sub1 = Subject 1, Sub2 = Subject 2.

Scrambled condition in which the long-term spectral features were conserved relative to the Natural Music condition but whose temporal features were effectively removed. Results from spectral and temporal control conditions show that the extent of ISS is greatly reduced for non-musical, compared with musical, stimuli in many of these brain regions. Most notably, sub-cortical auditory structures of the thalamus and midbrain also showed greater synchronization for the Natural Music condition. Additional analyses showed that the observed differences in ISS across stimulus conditions did not arise from stimulus-following, spectro-temporally invariant neural responses or synchronized movement, suggesting that the processing of music involves on-line cognitive and anticipatory processes and is not strictly stimulus-following (Huron, 2006). Taken together, our results indicate that a naturalistic and extended musical sequence elicits synchronized patterns of neural activity across individuals in auditory and motor regions of the brain as well as fronto-parietal regions associated with higher-level cognitive function, and that the structural content of a sound sequence is sufficient to dramatically alter synchronization throughout this extended network.

Sub-cortical synchronization to music

Our results show for the first time that sub-cortical structures of the auditory system are synchronized across subjects during music listening and include the IC of the midbrain and MGN of the thalamus bilaterally. IC is the primary midbrain nucleus in the auditory pathway, and auditory information processed in the IC is projected to auditory cortex via the MGN. Near-field (Creutzfeldt et al., 1980; Rees & Moller, 1983) and far-field (Griffiths et al., 2001) recordings from these sub-cortical auditory structures have shown that activity is driven by low-level acoustical features and a recent fMRI study showed orthogonal organization of spectral and temporal features in the primate IC (Baumann et al., 2011), corroborating evidence from near-field electrophysiological studies (Langner & Schreiner, 1988). Given that the temporal features in the Natural Music condition were effectively removed in the Phase-Scrambled condition, reduced ISS in sub-cortical (and cortical) structures for the Natural Music > Phase-Scrambled comparison was probably due to the fact that sub-cortical temporal processing mechanisms (Baumann et al., 2011) were weakly synchronized by the Phase-Scrambled stimulus while both spectral and temporal processing mechanisms were more strongly synchronized for the Natural Music condition. However, the interpretation for the Natural Music > Spectrally-Rotated result is different given that the Spectrally-Rotated condition contained the full complement of spectro-temporal features: the power spectrum was altered in this control condition but was not degraded or limited in any manner. Given the conservation of both temporal and spectral features in the Spectrally-Rotated condition, we hypothesize that the temporal structure of the Natural Music condition (Levitin & Menon, 2003; 2005) was responsible for the elevated ISS results in both sub-cortical and cortical regions relative to the control conditions.

These sub-cortical auditory structures have historically been considered passive relays of auditory information, and therefore it is surprising to find the strong enhancement in subcortical ISS in the Natural Music condition relative to the Spectrally-Rotated control condition. If these sub-cortical structures serve as passive relays of auditory information, then ISS should have been comparable for all stimulus conditions. In contrast to this hypothesis, our results indicate that ISS in sub-cortical structures is driven by the musical nature of the stimulus and suggest that top-down, cortically mediated influences play an important role in synchronizing activity in audi-

tory sub-cortical regions between subjects. This result is consistent with recent work showing that sub-cortical auditory structures are influenced by context (Chandrasekaran *et al.*, 2009), learning (Chandrasekaran *et al.*, 2012; Hornickel *et al.*, 2012; Skoe & Kraus, 2012; Anderson *et al.*, 2013) and memory (Tzounopoulos & Kraus, 2009).

An important question for all sub-cortical and cortical ISS findings is which aspect(s) of musical structure are responsible for the current ISS findings. Plausible candidates include themes, cadences, chord functions, tones, accents and dynamics, tempo, and any number of combinations of these features. The current work has controlled only for the contribution of spectro-temporal acoustical features to ISS and cannot provide additional information regarding musical features driving ISS results. An important avenue for future work is exploring the relative roles of these candidate musical features on ISS.

Synchronization in auditory structures of the temporal lobe

Our results demonstrate that auditory structures of the temporal lobe, including HG, PT, PP and pSTG bilaterally, were highly synchronized across subjects during music listening. Interestingly, no differences were evident in auditory cortical synchronization for the Natural Music > Spectrally-Rotated comparison, although differences were evident for the Natural Music > Phase-Scrambled comparison (Fig. 4). Amplitude modulation in the Natural Music and Spectrally-Rotated conditions is one possible explanation for ISS across both tasks in the auditory cortex. This interpretation is supported by previous studies which have shown auditory cortical sensitivity to low-frequency amplitude modulation in speech (Ahissar *et al.*, 2001; Abrams *et al.*, 2008, 2009; Aiken & Picton, 2008) and other auditory stimuli (Boemio *et al.*, 2005), and is further supported by single and multi-unit activity measured in auditory cortex of animal models during the processing of spectro-temporally complex auditory stimuli (Wang *et al.*, 1995; Nagarajan *et al.*, 2002). In this context it is noteworthy that a significant ISS difference was evident in auditory cortex for the Natural Music > Phase-Scrambled comparison (Fig. 4, right). These results indicate that despite the well-documented sensitivity of auditory cortex to spectral and harmonic information (Zatorre *et al.*, 2002), which are present in the Phase-Scrambled condition, these features alone, in the absence of temporal patterns, are insufficient to drive ISS. Our results extend these previous findings by showing that the disruption of temporal patterns in music significantly reduces the consistency of auditory cortical activity measured across individuals. Moreover, our results point to the involvement of both primary and secondary auditory cortical structures, including HG, PP, PT and pSTG, in tracking the temporal structure of music across time periods lasting minutes. Additionally, a recent ISS study showed that activity in bilateral STG and HG are recruited during timbral processing of a naturalistic musical stimulus, and bilateral STG and right-hemisphere HG are also active during rhythm processing (Alluri *et al.*, 2012). ISS results in the current study also support a role for STG and HG in rhythm processing given that (1) ISS in these auditory cortical regions was only evident when temporal features were present in the stimuli (see Fig. 4), and (2) temporal features, such as amplitude modulation, are fundamental to the perception of rhythm (Sethares, 2007).

An intriguing aspect of the results was the finding of differences in ISS for the Natural Music > Spectrally-Rotated condition in sub-cortical structures but not in auditory cortex. While both sub-cortical (Chandrasekaran *et al.*, 2009) and cortical structures (Fec-

teau *et al.*, 2004; Chait *et al.*, 2007) of the auditory system have shown sensitivity to the context of stimulus events, contextual processing is more closely associated with auditory cortex while stimulus-following is associated with sub-cortical structures. Very little is known about the relative influence of context on sub-cortical vs. cortical structures in the auditory system, and current models of the auditory system cannot easily explain this aspect of the results. It is hoped that future studies can address these questions further by examining functional interactions between multiple regions of the auditory hierarchy during the processing of extended stimulus sequences.

Synchronization in fronto-parietal cortex

An important new finding from our study is that ISS during music listening extends beyond auditory regions of superior temporal cortex. Of particular interest is the identification of right-lateralized regions of the IFG, including BAs 45 and 47, as well as the PGa subdivision of the inferior parietal cortex. Importantly, ISS was greater for the Natural Music condition compared with both control conditions in these fronto-parietal regions (Fig. 6). These brain structures have been implicated in previous studies of music processing: the IFG has been implicated in processing temporal structure (Levitin & Menon, 2003, 2005) and violations of syntactic structure (Maess *et al.*, 2001; Koelsch, 2005), and the AG has been implicated in musical memory (Platel *et al.*, 2003). Beyond the processing of these specific musical features, however, our results from the ISS analysis indicate that activity in these fronto-parietal structures is consistently synchronized to structural features in the musical stimulus, and suggest a role for these brain regions in the on-line tracking of musical structure. One possibility is that a fronto-parietal circuit involving right-hemisphere homologs of Broca's and Geschwind's areas support the processing of musical structure by engaging attentional and working memory resources necessary for the processing of extended nonlinguistic stimulus sequences. These resources are probably necessary for holding musical phrases and passages in mind as a means of tracking the long-term structure of a musical stimulus. Consistent with this hypothesis, a recent study examining expectation violation in response to brief string quartet compositions showed that right-hemisphere SMG and BA 44 of Broca's area are modulated by musical expertise, and may underlie enhanced attention and working memory function in musicians (Oechslin *et al.*, 2012).

Synchronization in motor planning regions of cortex

Our analysis also revealed significant ISS in the PMC, MCC and pre-central gyrus in response to the Natural Music condition, and ISS was greater in these brain regions for the Natural Music condition relative to the control conditions (Fig. 7B). The PMC and pre-central gyrus are associated with sensory-motor integration and motor imagery (Zatorre *et al.*, 2007; Sammler *et al.*, 2010). In a previous study it was shown that the PMC and pre-central gyrus are sensitive to the passive perception of musical rhythms, indicating that these regions can be activated in the absence of a motor task (Chen *et al.*, 2008). A plausible explanation of our results is that ISS in motor regions is driven by rhythmic components of the stimulus. Our study adds to this literature by showing that these motor planning regions are synchronized between subjects during a natural musical experience, and are likely time-locked to structural (e.g. rhythmic) components of the stimulus. One possible explanation for this connection with motor systems is that, over the course of

human evolution, music has traditionally been used in conjunction with synchronized movement and dance (McNeill, 1995; Levitin, 2008).

Conclusion

Our study provides new information regarding inter-subject brain synchronization in response to natural stimuli. Our results show that inter-subject synchronization occurs at multiple levels in the information processing hierarchy – from sub-cortical and cortical auditory structures to fronto-parietal attention network and motor planning areas. Importantly, we show for the first time that this diverse collection of auditory and supra-auditory brain structures tracks aspects of musical structure over extended periods of time. More generally, our findings demonstrate that music listening elicits consistent and reliable patterns of time-locked brain activity in response to naturalistic stimuli that extends well beyond primary sensory cortices (Hasson *et al.*, 2004; Wilson *et al.*, 2008), and that synchronization is not driven solely by low-level acoustical cues. These signatures of synchronized brain activity across individuals in multiple hierarchically structured systems may underlie shared neural representations that facilitate our collective social capacity for listening and attending to music.

Supporting Information

Additional supporting information can be found in the online version of this article:

Fig. S1. Differences between ISS and GLM approaches for the analysis of music processing in the brain.

Fig. S2. Flow chart for ISS Analysis. Synchronization was calculated by computing Pearson correlations between the voxel time series in each pair of subjects (136 subject-to-subject comparisons total).

Data S1. Methods.

Acknowledgements

Abbreviations

AG, angular gyrus; fMRI, functional magnetic resonance imaging; GLM, general linear model; HG, Heschl's gyrus; IC, inferior colliculus; IFG, inferior frontal gyrus; IPS, intra-parietal sulcus; ISS, inter-subject synchronization; MCC, mid-cingulate cortex; MGN, medial geniculate nucleus; PGa and PGp, anterior and posterior sub-divisions of the angular gyrus; PMC, premotor cortex; PP, planum polare; pSMG, posterior supramarginal gyrus; pSTG, posterior superior temporal gyrus; PT, planum temporale.

References

Abrams, D.A., Nicol, T., Zecker, S. & Kraus, N. (2008) Right-hemisphere auditory cortex is dominant for coding syllable patterns in speech. *J. Neurosci.*, **28**, 3958–3965.

Abrams, D.A., Nicol, T., Zecker, S. & Kraus, N. (2009) Abnormal cortical processing of the syllable rate of speech in poor readers. *J. Neurosci.*, **29**, 7686–7693.

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. *Cereb. Cortex*, **21**, 1507–1518.

Abrams, D.A., Ryali, S., Chen, T., Balaban, E., Levitin, D.J. & Menon, V. (2012) Multivariate activation and connectivity patterns discriminate speech intelligibility in Wernicke's, Broca's, and Geschwind's Areas. *Cereb. Cortex*, doi: 10.1093/cercor/bhs165 [Epub ahead of print].

Ahissar, E., Nagarajan, S., Ahissar, M., Protopapas, A., Mahncke, H. & Merzenich, M.M. (2001) Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proc. Natl. Acad. Sci. USA*, **98**, 13367–13372.

Aiken, S.J. & Picton, T.W. (2008) Human cortical responses to the speech envelope. *Ear Hearing*, **29**, 139–157.

Alluri, V., Toiviainen, P., Jaaskelainen, I.P., Glerean, E., Sams, M. & Brattico, E. (2012) Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm. *NeuroImage*, **59**, 3677–3689.

Anderson, S., White-Schwoch, T., Parbery-Clark, A. & Kraus, N. (2013) Reversal of age-related neural timing delays with training. *Proc. Natl. Acad. Sci. USA*, **110**, 4357–4362.

Baumann, S., Griffiths, T.D., Sun, L., Petkov, C.I., Thiele, A. & Rees, A. (2011) Orthogonal representation of sound dimensions in the primate midbrain. *Nat. Neurosci.*, **14**, 423–425.

Blesser, B. (1972) Speech perception under conditions of spectral transformation. 1. Phonetic characteristics. *J. Speech Hear. Res.*, **15**, 5–41.

Boemio, A., Fromm, S., Braun, A. & Poeppel, D. (2005) Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nat. Neurosci.*, **8**, 389–395.

Caspers, S., Geyer, S., Schleicher, A., Mohlberg, H., Amunts, K. & Zilles, K. (2006) The human inferior parietal cortex: cytoarchitectonic parcellation and interindividual variability. *NeuroImage*, **33**, 430–448.

Chait, M., Poeppel, D. & Simon, J.Z. (2007) Stimulus context affects auditory cortical responses to changes in interaural correlation. *J. Neurophysiol.*, **98**, 224–231.

Chandrasekaran, B., Hornickel, J., Skoe, E., Nicol, T. & Kraus, N. (2009) Context-dependent encoding in the human auditory brainstem relates to hearing speech in noise: implications for developmental dyslexia. *Neuron*, **64**, 311–319.

Chandrasekaran, B., Kraus, N. & Wong, P.C.M. (2012) Human inferior colliculus activity relates to individual differences in spoken language learning. *J. Neurophysiol.*, **107**, 1325–1336.

Chen, J.L., Penhune, V.B. & Zatorre, R.J. (2008) Listening to musical rhythms recruits motor regions of the brain. *Cereb. Cortex*, **18**, 2844–2854.

Creutzfeldt, O., Hellweg, F.C. & Schreiner, C. (1980) Thalamocortical transformation of responses to complex auditory stimuli. *Exp. Brain Res.*, **39**, 87–104.

Deike, S., Gaschler-Markefski, B., Brechmann, A. & Scheich, H. (2004) Auditory stream segregation relying on timbre involves left auditory cortex. *NeuroReport*, **15**, 1511–1514.

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.

Glover, G.H. & Lai, S. (1998) Self-navigated spiral fMRI: interleaved versus single-shot. *Magn. Reson. Med.*, **39**, 361–368.

Grahn, J.A. & Rowe, J.B. (2009) Feeling the beat: premotor and striatal interactions in musicians and nonmusicians during beat perception. *J. Neurosci.*, **29**, 7540–7548.

Griffiths, T.D., Uppenkamp, S., Johnsrude, I., Josephs, O. & Patterson, R.D. (2001) Encoding of the temporal regularity of sound in the human brainstem. *Nat. Neurosci.*, **4**, 633–637.

Grinsted, A., Moore, J.C. & Jevrejeva, S. (2004) Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Proc. Geoph.*, **11**, 561–566.

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G. & Malach, R. (2004) Intersubject synchronization of cortical activity during natural vision. *Science*, **303**, 1634–1640.

Hasson, U., Malach, R. & Heeger, D.J. (2010) Reliability of cortical activity during natural stimulation. *Trends Cogn. Sci.*, **14**, 40–48.

Hornickel, J., Zecker, S.G., Bradlow, A.R. & Kraus, N. (2012) Assistive listening devices drive neuroplasticity in children with dyslexia. *Proc. Natl. Acad. Sci. USA*, **109**, 16731–16736.

Huron, D. (2006) *Sweet Anticipation: Music and the Psychology of Expectation*. MIT Press, Cambridge, MA.

Janata, P., Birk, J.L., Van Horn, J.D., Leman, M., Tillmann, B. & Bharucha, J.J. (2002) The cortical topography of tonal structures underlying Western music. *Science*, **298**, 2167–2170.

Joris, P.X., Schreiner, C.E. & Rees, A. (2004) Neural processing of amplitude-modulated sounds. *Physiol. Rev.*, **84**, 541–577.

Kaas, J.H. & Hackett, T.A. (2000) Subdivisions of auditory cortex and processing streams in primates. *Proc. Natl. Acad. Sci. USA*, **97**, 11793–11799.

Koelsch, S. (2005) Neural substrates of processing syntax and semantics in music. *Curr. Opin. Neurobiol.*, **15**, 207–212.

Langner, G. & Schreiner, C.E. (1988) Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. *J. Neurophysiol.*, **60**, 1799–1822.

Leaver, A.M., Van Lare, J., Zielinski, B., Halpern, A.R. & Rauschecker, J.P. (2009) Brain activation during anticipation of sound sequences. *J. Neurosci.*, **29**, 2477–2485.

Levitin, D.J. (2008) *The World in Six Songs: How the Musical Brain Created Human Nature*. Dutton/Penguin, New York.

Levitin, D.J. & Menon, V. (2003) Musical structure is processed in 'language' areas of the brain: a possible role for Brodmann Area 47 in temporal coherence. *NeuroImage*, **20**, 2142–2152.

Levitin, D.J. & Menon, V. (2005) The neural locus of temporal structure and expectancies in music: evidence from functional neuroimaging at 3 Tesla. *Music Percept.*, **22**, 563–575.

Maess, B., Koelsch, S., Gunter, T.C. & Friederici, A.D. (2001) Musical syntax is processed in Broca's area: an MEG study. *Nat. Neurosci.*, **4**, 540–545.

McNeill, W.H. (1995) *Keeping Togther in Time: Dance and Drill in Human History*. Harvard University Press, Cambridge, MA.

Menon, V., Levitin, D.J., Smith, B.K., Lembke, A., Krasnow, B.D., Glazer, D., Glover, G.H. & McAdams, S. (2002) Neural correlates of timbre change in harmonic sounds. *NeuroImage*, **17**, 1742–1754.

Muhlau, M., Rauschecker, J.P., Oestreicher, E., Gaser, C., Rottinger, M., Wohlschlager, A.M., Simon, F., Etgen, T., Conrad, B. & Sander, D. (2006) Structural brain changes in tinnitus. *Cereb. Cortex*, **16**, 1283–1288.

Nagarajan, S.S., Cheung, S.W., Bedenbaugh, P., Beitel, R.E., Schreiner, C.E. & Merzenich, M.M. (2002) Representation of spectral and temporal envelope of twitter vocalizations in common marmoset primary auditory cortex. *J. Neurophysiol.*, **87**, 1723–1737.

Nichols, T.E. & Holmes, A.P. (2002) Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum. Brain Mapp.*, **15**, 1–25.

Oechslin, M.S., Van De Ville, D., Lazeyras, F., Hauert, C.A. & James, C.E. (2012) Degree of musical expertise modulates higher order brain functioning. *Cereb. Cortex*, doi: 10.1093/cercor/bhs206 [Epub ahead of print].

Passynkova, N., Sander, K. & Scheich, H. (2005) Left auditory cortex specialization for vertical harmonic structure of chords. *Ann. NY Acad. Sci.*, **1060**, 454–456.

Patel, A.D. & Balaban, E. (2001) Human pitch perception is reflected in the timing of stimulus-related cortical activity. *Nat. Neurosci.*, **4**, 839–844.

Platel, H., Baron, J.C., Desgranges, B., Bernard, F. & Eustache, F. (2003) Semantic and episodic memory of music are subserved by distinct neural networks. *NeuroImage*, **20**, 244–256.

Prichard, D. & Theiler, J. (1994) Generating surrogate data for time-series with several simultaneously measured variables. *Phys. Rev. Lett.*, **73**, 951–954.

Rees, A. & Moller, A.R. (1983) Responses of neurons in the inferior colliculus of the rat to AM and FM tones. *Hearing Res.*, **10**, 301–330.

Rorden, C. & Brett, M. (2000) Stereotaxic display of brain lesions. *Behav. Neurol.*, **12**, 191–200.

Sammler, D., Baird, A., Valabregue, R., Clement, S., Dupont, S., Belin, P. & Samson, S. (2010) The relationship of lyrics and tunes in the processing of unfamiliar songs: a functional magnetic resonance adaptation study. *J. Neurosci.*, **30**, 3572–3578.

Scott, S.K., Blank, C.C., Rosen, S. & Wise, R.J.S. (2000) Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*, **123**, 2400–2406.

Sethares, W.A. (2007) *Rhythm and Transforms*. Springer, London.

Skoe, E. & Kraus, N. (2012) A little goes a long way: how the adult brain is shaped by musical training in childhood. *J. Neurosci.*, **32**, 11507–11510.

Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., Niazy, R.K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J.M. & Matthews, P.M. (2004) Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, **23** (Suppl. 1), S208–S219.

Snyder, J.S. & Large, E.W. (2005) Gamma-band activity reflects the metric structure of rhythmic tone sequences. *Brain Res. Cogn. Brain Res.*, **24**, 117–126.

Song, J.H., Skoe, E., Wong, P.C. & Kraus, N. (2008) Plasticity in the adult human auditory brainstem following short-term linguistic training. *J. Cognitive Neurosci.*, **20**, 1892–1902.

Sridharan, D., Levitin, D.J., Chafe, C.H., Berger, J. & Menon, V. (2007) Neural dynamics of event segmentation in music: converging evidence for dissociable ventral and dorsal networks. *Neuron*, **55**, 521–532.

Torrence, C. & Compo, G. (1998) A practical guide to wavelet analysis. *B. Am. Meteorol. Soc.*, **79**, 61–78.

Tzounopoulos, T. & Kraus, N. (2009) Learning to encode timing: mechanisms of plasticity in the auditory brainstem. *Neuron*, **62**, 463–469.

Wang, X., Merzenich, M.M., Beitel, R. & Schreiner, C.E. (1995) Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: temporal and spectral characteristics. *J. Neurophysiol.*, **74**, 2685–2706.

Warren, J.E., Sauter, D.A., Eisner, F., Wiland, J., Dresner, M.A., Wise, R.J., Rosen, S. & Scott, S.K. (2006) Positive emotions preferentially engage an auditory-motor 'mirror' system. *J. Neurosci.*, **26**, 13067–13075.

Wilson, S.M., Molnar-Szakacs, I. & Iacoboni, M. (2008) Beyond superior temporal cortex: intersubject correlations in narrative speech comprehension. *Cereb. Cortex*, **18**, 230–242.

Zatorre, R.J., Evans, A.C. & Meyer, E. (1994) Neural mechanisms underlying melodic perception and memory for pitch. *J. Neurosci.*, **14**, 1908–1919.

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

Zatorre, R.J., Chen, J.L. & Penhune, V.B. (2007) When the brain plays music: auditory-motor interactions in music perception and production. *Nat. Rev. Neurosci.*, **8**, 547–558.