OXFORD

Temporal dynamics of musical emotions examined through intersubject synchrony of brain activity

Wiebke Trost,[1,2] Sascha Frühholz,[1,2] Tom Cochrane,[3] Yann Cojan,[1,4] and Patrik Vuilleumier[1,4]

[1]Swiss Center of Affective Sciences, University of Geneva, Geneva, Switzerland, [2]Neuroscience of Emotions and Affective Dynamics Laboratory, University of Geneva, Geneva, Switzerland, [3]Department of Philosophy, University of Sheffield, Sheffield, UK, and [4]Laboratory for Behavioral Neurology and Imaging of Cognition, University of Geneva, Geneva, Switzerland

Correspondence should be addressed to Wiebke J. Trost, Swiss Center of Affective Sciences, Biotech Campus, University of Geneva, 9, Chemin des Mines, 1202 Geneva, Switzerland. E-mail: Johanna.trost@unige.ch

Abstract

To study emotional reactions to music, it is important to consider the temporal dynamics of both affective responses and underlying brain activity. Here, we investigated emotions induced by music using functional magnetic resonance imaging (fMRI) with a data-driven approach based on intersubject correlations (ISC). This method allowed us to identify moments in the music that produced similar brain activity (i.e. synchrony) among listeners under relatively natural listening conditions. Continuous ratings of subjective pleasantness and arousal elicited by the music were also obtained for the music outside of the scanner. Our results reveal synchronous activations in left amygdala, left insula and right caudate nucleus that were associated with higher arousal, whereas positive valence ratings correlated with decreases in amygdala and caudate activity. Additional analyses showed that synchronous amygdala responses were driven by energy-related features in the music such as root mean square and dissonance, while synchrony in insula was additionally sensitive to acoustic event density. Intersubject synchrony also occurred in the left nucleus accumbens, a region critically implicated in reward processing. Our study demonstrates the feasibility and usefulness of an approach based on ISC to explore the temporal dynamics of music perception and emotion in naturalistic conditions.

Key words: intersubject correlation; musical emotion; musical features; amygdala; fMRI

Introduction

The power of music to arouse strong emotions in listeners is rooted in its essentially temporal nature. This quality enables music to capture the nuance and flux of emotional experience in a way that listeners can empathically track from moment to moment (Cochrane, 2010). It is therefore crucial to consider the temporal dynamics of music processing in order to better understand the generation of emotional reactions and uncover the responsible brain mechanisms. However, music represents a very complex acoustic stimulus, whose impact on the listener is challenging to characterize. Although it has been observed that musical works are capable of 'striking a common chord'; evoking consistent emotions across different individuals (Panksepp and Bernatzky, 2002; Guhn *et al.*, 2007), it remains difficult to predict which kind of musical event or which acoustic features will trigger a specific emotion. In this study, we tackle this problem by probing the underlying neural activities that listeners share while listening to musical works. Specifically, we investigated the relationship between the two major properties of music, namely, its emotional and temporal dimensions. More generally, valuable insight into music perception can be gained by better understanding this relationship and pinpointing their substrates in the human brain.

Several neuroimaging studies have investigated emotions induced by music (Blood and Zatorre, 2001; Menon and Levitin, 2005; Salimpoor *et al.*, 2011; Trost *et al.*, 2012; Koelsch, 2014). However, these studies typically used block designs averaging brain responses to different kinds of musical excerpts over a period of several seconds or minutes, and were thus generally based on strong *a priori* assumptions about the emotional impact and homogeneity of a whole musical excerpt. Hence, even though this research has yielded many precious insights (Koelsch, 2014), it did not take into account the temporal dynamics of how musical emotions may develop in time. Yet, in their very essence, emotions are defined as transient changes in mental and physiological states (Scherer, 2004). Therefore, our study represents a significant advance on these earlier studies by focusing on transient fluctuations in brain activity evoked by musical stimuli in a consistent manner across different individuals. To this aim, we applied a novel data-driven methodological approach based on intersubject correlations (ISC).

The method of ISCs to investigate brain activity with functional magnetic resonance imaging (fMRI) has first been introduced by Hasson *et al.* (2004) to study visual perception during movie viewing under naturalistic conditions. By calculating pairwise correlations of the fMRI signal of different participants and then averaging the results at the group level, this technique permits the identification of brain areas that show a similar time course of activation across subjects, without the need to define events *a priori*. Further, by using reverse-correlation as a mean to trace back in the original stimulus (e.g. movie) when high ISC occurred for specific brain regions (corresponding to moments of increased 'synchrony' between individuals), Hasson *et al.* (2004) demonstrated that scenes with close-ups on faces were responsible for increased ISC in the fusiform cortex, a visual area critically implicated in face recognition (Kanwisher *et al.*, 1997). Conversely, viewing hand movements accounted for higher ISC in somatosensory and motor cortical areas. Several other studies have also used ISC analysis as a data-driven approach to study the attentional and emotional effects of films (Furman *et al.*, 2007; Hasson *et al.*, 2008a,b, 2010; Jaaskelainen *et al.*, 2008), as well as speech processing (Wilson *et al.*, 2008) and more recently music perception (Alluri *et al.*, 2011; Abrams *et al.*, 2013). But these studies used this methodology mostly to determine common fluctuations in brain activity at different frequency ranges or to identify large-scale networks recruited in these task conditions in general. For example, Abrams *et al.* (2013) compared natural symphonic music with scrambled versions of the same piece in order to identify brain regions that would specifically synchronize their activity during natural music. These authors found that auditory cortices, thalamus, motor areas, as well as frontal and parietal regions synchronize more across participants when listening to natural music than to scrambled versions. However, beyond synchronization across entire musical pieces, ISC can be applied and be particularly informative over smaller time scales in order to identify specific moments at which the sensory inputs will yield high intersubject agreement in neural activity. Thus, as successfully used by Hasson *et al.* (2004), this approach combined with reverse correlation provides a very powerful tool to unveil unique events in music pieces that generate similar brain responses and consistent emotional reactions in the listeners.

In this study, we employed a novel ISC approach to identify moments in time during which similar changes in brain activity arose while people listened to the same piece of music, without any *a priori* definition of these moments from music content

itself. As music can elicit powerful and consistent emotional reactions (Scherer, 2004; Juslin, 2013), we predicted that emotional experiences shared by different people at the same moment of the music should correspond to the recruitment of similar brain processes, and thus lead to enhanced synchrony of activation patterns between listeners. Furthermore, if ISC occurs at particular moments in the music, we also predicted that it may reflect the occurrence of specific features in the music (e.g. in terms of acoustic content or subjective affect) and tested which features drove ISC in different brain regions. This approach differs from the standard GLM approach based on a model of the BOLD response to predefined stimuli in the sense that here no preclassified stimuli or conditions are chosen a priori to determine the corresponding brain responses. Instead, moments of coherent brain activity across participants were first determined and subsequently used to analyze whether (and which) aspects of emotional feelings or acoustical inputs triggered these synchronized brain responses. In addition, unlike a traditional GLM analysis based on the time course of acoustic features or emotion ratings, our measure of ISC may potentially detect musical events eliciting particular brain responses even when they are not strictly simultaneous with an increase in specific music features or show non-linear relationships with these features (e.g. some ISC and corresponding emotional event may arise only after reaching a certain level).

Previous neuroimaging studies have already determined brain regions involved in the processing of various acoustic features for speech, prosody and music (Platel *et al.*, 1997; Zatorre *et al.*, 2002; Schirmer and Kotz, 2006; Alluri *et al.*, 2011, Fruhholz *et al.*, 2012). A few of these studies used correlation approaches that directly regress brain activity from each participant with particular characteristics of the stimulus (e.g. loudness or dissonance) over its whole presentation period (Chapin *et al.*, 2010; Alluri *et al.*, 2011). This work has established that elementary acoustic features are processed in primary auditory areas, then integrated into a coherent percept in secondary auditory regions, and finally categorized as a function of cognitive and affective meaning in higher level cortical areas such as the inferior frontal gyrus and orbitofrontal cortex, plus cortical and subcortical motor structures (Schirmer and Kotz, 2006, Koelsch, 2011). Meanwhile, from psychological studies, it is also known that certain acoustic features of music are linked with particular emotional responses (Gabrielson and Juslin, 2003). For example, fast tempo and high volume both induce higher subjective arousal (Schubert, 2004; Etzel *et al.*, 2006; Gomez and Danuser, 2007), whereas dissonance is closely related to perceived (un)pleasantness (Blood *et al.*, 1999; Koelsch *et al.*, 2006). Dissonance has also been associated with evocations of excitement, fear and anger, as well as with sadness (Krumhansl, 1997). Timbre sharpness correlates with perception of anger and timbre brightness with happiness (Juslin and Laukka, 2000), whereas a flat spectrum with many harmonics has been related to disgust and fear (Scherer and Oshinsky, 1977). Finally, rhythmic stability is linked with relaxation, peacefulness or happiness (Hevner, 1936). Here, by determining acoustic features and emotional responses concomitant to increased synchrony of brain activity between listeners, we were able to gain better knowledge about both the auditory characteristic and neural structures underlying music-induced emotions.

Although the exact music dimensions responsible for differentiated affective experiences remain poorly known (Juslin, 2013; Koelsch, 2013), the involvement of distributed networks of sensory, motor and limbic regions showing distinctive patterns of activation according to type of emotion elicited by the music

have been observed (Trost *et al.*, 2012; Koelsch and Skouras, 2014). The ventral striatum, but also ACC, insula and rolandic parietal operculum are generally recruited by pleasant emotions induced by music such as joy or wonder (Blood and Zatorre, 2001; Koelsch *et al.*, 2006; Mitterschiffthaler *et al.*, 2007; Trost *et al.*, 2012). On the other hand, negative musical emotions such as sadness, fear or tension, have been associated with the amygdala, hippocampus and ventral ACC (Koelsch *et al.*, 2006; Green *et al.*, 2008). Amygdala activity also correlates negatively with the intensity of perceived chills to one's favorite music (Blood and Zatorre, 2001).

While these studies represent significant progress on the neural bases of musical emotions, they mostly focused on emotional experiences reported for the whole duration of a music piece, rather than more transient emotions evoked at precise time points during listening. In contrast our study allowed us to investigate emotions induced by music at different moments in a piece using an ecological design during relatively natural listening conditions. Participants heard a full movement from a work in the classical repertoire (9–14 min) while undergoing fMRI. Using a data-driven approach based on ISC, we first identified those moments in the music where common brain responses were evoked across different listeners. These peaks in ISC were then analyzed in terms of the musical features present at the same moment and the corresponding emotional ratings made by listeners using a continuous dynamic judgment procedure. We predicted that distinct brain regions, including not only sensory cortices, but also limbic structures and basal ganglia, would differentially be recruited by different emotion dimensions and respond to different music features. Note that our method aimed at identifying common and synchronized brain responses across all listeners, but not more idiosyncratic reactions of individual participants that did not occur in other participants at the same moment. Importantly, however, this novel approach shares the virtue of avoiding *a priori* assumptions about specific dimensions of the stimuli, as used by other studies (Hasson *et al.*, 2004, 2010) to investigate vision. Thus, by extracting meaningful information from ongoing brain activity in naturalistic conditions, measures of ISC provide a powerful data-driven methodology that may offer valuable information about music processing and music-induced emotions.

Method

Subjects

Seventeen healthy participants (mean age: 25.4 years; s.d. ± 5.1, 9 females) with no history of neurological or psychiatric disease took part in the fMRI experiment. Another 14 participants (mean age: 28.7 years; s.d. ± 6.6, 8 females) took part in the behavioral experiment. All participants reported normal hearing, a liking for classical music, and had a minimum of 5 years of practical musical training. Participants in the fMRI experiment were all right handed. All subjects were paid for their participation and gave informed consent in accord with the regulation of the local ethics committee. The study was approved by the local ethics committee.

Stimulus material

Musical stimuli consisted of three complete movements from the classical music repertoire (see Supplementary Table S1). The entire movements were presented to study music listening under ecologically valid conditions. To avoid the potential influence of lyrics, only instrumental pieces were chosen. The duration of each piece was between 9 and 14 min. All pieces were in the 'theme and variations' form, which allowed segmenting the music in a well-defined manner. The number of variations varied between the pieces (Mendelssohn: 18 segments, Prokofiev: 8 segments, Schubert: 14 segments). Furthermore the pieces differed a great deal in the style of composition, the composer and instrumentation. These differences were chosen in order to avoid a biased measure focusing only on one particular musical style and thereby allowing more reliable inferences to be drawn about general emotional qualities shared by these works

Experimental design of fMRI experiment

For the fMRI experiment, participants were instructed and familiarized with the task prior to scanning. They had to listen attentively to the music throughout the presentation of each piece, while keeping their eyes open and directed toward the middle of a dimly-lit screen. During this time, they were asked to let their emotions naturally flow with the music, while constraining overt bodily movements. At the end of every musical stimulus, participants answered a short questionnaire asking for their global subjective evaluation of the previously heard music. A series of six questions were presented on the screen, one after the other, immediately after the end of the music excerpt. Participants had to rate (i) how much attention they paid to the piece, (ii) how strongly they reacted emotionally and (iii) how much the environment influenced them throughout the presentation of the previously presented piece. In addition they also had to rate (iv) how much they liked the piece and (v) how familiar they were with it. Finally, for each piece, the participants had to evaluate their subjective emotional experience along the nine GEMS dimensions of the Geneva Emotion Music Scale (Zentner *et al.*, 2008), including wonder, joy, power, tenderness, peacefulness, nostalgia, tension, sadness and transcendence. All evaluations ('I felt this emotion/I agree') had to be made on a horizontal seven-point scale ranging from 0 (not at all) to 6 (absolutely). During the presentation of the questionnaires, image acquisition by the scanner stopped and there was no time limit for answering. Answering all questions took less than 2 min on average.

During music listening, the screen was grey and luminosity inside the scanner room was kept low and constant. Visual instructions were presented on a screen back-projected on a headcoil-mounted mirror. Musical stimuli were presented binaurally with a high-quality MRI-compatible headphone system (CONFON HP-SC 01 and DAP-center mkII, MR confon GmbH, Germany). Participants used a MRI compatible response button box (HH-1×4-CR, Current Designs Inc., USA) to give their answers. The fMRI experiment consisted of three scanning runs and ended with the acquisition of the structural MRI scan. The presentation order of the three musical pieces was randomized between participants. Prior to scanning, the loudness of the music was adjusted for each participant individually to ensure that both the loudest and softest moments of the musical works could be comfortably heard.

fMRI data acquisition and preprocessing

MRI images were acquired using a 3T whole body MRI scanner (Trio TIM, Siemens, Germany) together with a 12-channel head coil. A high-resolution T1-weighted structural image $(1 \times 1 \times 1\,mm^3)$ was obtained using a magnetization-prepared

rapid acquisition gradient echo sequence [repetition time (TR) = 1.9 s, echo time (TE) = 2.27 ms, time to inversion (TI) = 900 ms]. Functional images were obtained using the following parameters: 35 slices, slice thickness 3.2 mm, slice gap 20%, TR = 2 s, TE = 30 ms, field of view = 205 × 205 mm^2, 64 × 64 matrix, flip angle: 80°. FMRI data was analyzed using Statistical Parametric Mapping (SPM8; Wellcome Trust Center for Imaging, London, UK; http://www.fil.ion.ucl.ac.uk/spm). Data processing included realignment, unwarping, slice timing, normalization to the Montreal Neurological Institute space using an EPI template (resampling voxel size: 2 × 2 × 2 mm) and spatial smoothing (5 mm full-width at half-maximum Gaussian Filter).

fMRI data analysis

ISCs in the time-course of the preprocessed data were calculated using the ISCs toolbox (isc-toolbox, release 1.1) developed by Kauppi *et al.* (2010, 2014). Each piece was first analyzed separately over its entire time course to obtain global ISC maps of brain areas generally activated during music listening. Dynamic ISCs were also calculated using a sliding time windows with a bin size of 10 scans and a step size of 1 scan. Significance thresholds were determined by non-parametric permutation tests for the global maps and for the time-window analysis for each piece individually. The dynamic time-window based ISCs were used to identify peak moments of high intersubject agreement in fMRI activity over the whole brain. These peaks were defined such that at least 10 voxels or 10% of the voxels in a specific region of interest (ROI) would survive the calculated statistical threshold for the time-window ISC. Subsequently, the magnitude of fMRI activity during the corresponding time-window was extracted from these voxels, averaged over all participants and submitted to further correlation analyses with the emotional ratings and with the acoustic musical features identified for the corresponding time-window (second-order correlation). Given a window size of 10 scans, 10 time points were thus taken into account for the correlation. For these correlation analyses, a delay of 4 s was assumed in order to take into account the delay of the BOLD signal. For each ROI, the obtained correlation coefficients were Fishers-z-transformed and tested for significance using one-sample *t*-tests. The anatomical definition of ROIs was based on the HarvardOxford atlas (http://www.cma.mgh.harvard.edu/fsl_atlas.html). We selected 29 regions including auditory, motor and prefrontal areas, as well as subcortical regions, according to our main hypotheses about music perception and emotional processing (see Supplementary Tables S4 and S5 for the complete list of ROIs). For all ROIs, the left and right side were analyzed separately. ROIs were kept for further analyses only when there were more than five time windows that survived the statistical threshold across the three pieces.

Musical feature analysis

The three music pieces were analyzed in terms of acoustic features using the MIR toolbox (Lartillot and Toiviainen, 2007) implemented in Matlab. We selected different acoustic features describing several timbral and rhythmic characteristics of music (see Supplementary Table S2). These features included two long-term characteristics (pulse clarity and event density) that capture structural and context-dependent aspects of rhythmic variations (Lartillot *et al.*, 2008), as well as six short-term characteristics that concern loudness, timbre and frequency content (root mean square (rms), dissonance, zero-crossing, spectral entropy, spectral centroid, brightness). The loudness of the music is determined by the rms of the signal. Zero-crossing and entropy are measures indicating the noisiness of the acoustic signal, whereas centroid describes the central tendency of the spectrum of the signal. Brightness represents the amount of high-frequency energy contained in the signal. For long-term features, the music was analyzed in time-windows of 3 s that were shifted in 1 s steps. Short-term features were measured in successive windows of 50 ms with shifts of 25 ms. These parameters were chosen in accordance with previous publications using a similar methodology for musical feature analysis (Alluri *et al.*, 2011). In addition, the absolute tempo of pieces was extracted using built-in functions of the software Melodyne (version 3.2, Celemony Software, GmbH, München, Germany, www.celemony.com) refined to capture beat-by-beat tempo fluctuations. All musical features were consequently downsampled to 0.5 Hz. Given that some features were correlated with each other, a factor analysis was conducted in order to compare the different types of features and identify possible higher-order factors corresponding to coherent categories of inter-related features.

Experimental design of behavioral experiment

The behavioral study was conducted in a similar manner as the fMRI experiment, using identical musical stimuli as for the fMRI experiment. The main difference was that instead of passive listening, the participants had now to perform an active continuous evaluation task. The task consisted of dynamic judgments of subjectively felt emotions in response to the music over time. Participants were instructed to use an in-lab programmed interface (Adobe Flash CS3 Professional, ActionScript 3 in Flash, http://www.adobe.com/fr/), which allowed them to indicate the level of their affective experience by moving the computer mouse vertically along an intensity axis (see Supplementary Figure S2). The evaluation was displayed continuously on the computer screen. The intensity responses were recorded at a frequency of 4 Hz.

The experiment was conducted in two different versions. One group of participants (n = 14) was asked to evaluate their subjective affective state along the arousal dimension, going from 'calm-sleepy' to 'aroused-excited'. A second group of participants (n = 17) had to evaluate their affective state along the valence dimension, going from 'unpleasant-negative' to 'pleasant-positive'. Those participants who participated in the fMRI experiment were re-contacted and asked to perform the valence rating task 3 months after scanning, whereas a new group of different participants was invited to perform the arousal rating task. Both groups of participants were instructed to listen attentively to the music at the same time as they introspectively monitored and judged the felt level of valence or arousal evoked in them by the music. It was explicitly pointed out that they should not evaluate what they thought that the compositor or musicians wanted to express. In addition, as in the fMRI experiment, each musical piece was immediately followed by a short questionnaire of the same six questions concerning their global feelings as experienced overall during the whole duration of the piece. Participants in the behavioral experiment were comfortably installed in front of the computer and the musical stimuli were presented via high-quality headphones (Sennheiser, HD280 pro). Ratings were made by moving (with the right hand) the mouse up or down relative to a neutral baseline presented on the screen. Note that obtaining valence and arousal ratings from different groups allowed us to take

into account individual variability in the subjective pleasantness of the different music pieces, while minimizing habituation in arousal measures due to repeated exposure to the same stimuli. Presenting the same pieces three times for different ratings to the same group of participants would be likely to bias the data by changing the emotional effects produced by the music when it is heard several times.

The entire duration of the experiment took around 45 min. There was a training trial first, to adjust the volume and get familiarized with the task and the interface. Participants could repeat the training several times if necessary. We used the terms 'energy' and 'alertness' as definition of the arousal dimension in order to be clearer for naive participants and ensure that they understood well the nature of this dimension (see Supplementary Figure S2).

Behavioral data analysis

To analyze the rating data of this experiment, the continuous time courses of valence and arousal values were z-transformed and averaged in each of the two groups. To assess the variability of evaluations between participants, we calculated ISC between all pairs of subjects and averaged per subject the correlation coefficients with all other participants, after transforming the coefficients using the Fishers z-transformation.

Data from the short questionnaires at the end of each block were used to compare whether the emotional perception of the musical pieces was similar between groups in the behavioral and fMRI experiments, and whether the emotional experience of the fMRI group would be comparable after 3 months. We used a mixed ANOVA with *context* (inside/outside of scanner) as categorical variable and depending on the analysis *piece* or *GEMS* as single within factor. These statistical analyses were done in order to test if all the groups of the participants did the experiment in similar conditions and had similar emotional reactions to the music.

Results

Behavioral results

The global behavioral ratings given at the end of each music piece were compared between the different groups of participants and the different experimental contexts (inside or outside the scanner). Participants in both context conditions reported a similar amount of emotional reactions to the presentation of the three pieces and reported similar familiarity with these pieces (Table 1a, Supplementary Figure S1). However, participants in the scanner reported to be more distracted by the environment and to be slightly less attentive to the music than the participants outside of the scanner, as could be expected due to the MRI environment and noise. However, inside the scanner, participants evaluated all pieces as more pleasant than the participants outside. This shows that despite the distraction due to the scanner, participants were able to appreciate the music. There were no differences between the pieces concerning the attentional level and environmental distraction. However, the amount of the emotional reactions, as well as ratings of pleasantness and familiarity differed between the pieces. The Schubert piece was evaluated as more pleasant, more familiar and elicited more emotional reactions than either of the other pieces (Table 1a, Supplementary Figure S1). None of these ratings showed any significant interaction between the context condition and the musical pieces (Table 1a).

Likewise, emotion ratings along the nine categories of the GEMS (Zentner *et al.*, 2008) did not vary between the different context conditions for the pieces of Prokofiev and Schubert, but there was an interaction of context by category for the piece of Mendelssohn reflecting the fact that participants in the fMRI context evaluated this piece as conveying stronger feelings of 'power' relative to other contexts (Table 1b, Supplementary Figure S1). Because this piece ends with a powerful finale, this might have been produced by a recency effect that was attenuated in behavioral experiment conditions where participants had already provided ratings on other dimensions for the duration of the work.

The continuous judgments were analyzed for consistency across participants for each music piece (Table 2, Figure 1). For arousal, participants evaluated the pieces very similarly (average Cronbach alpha = 0.967). For valence, there was a greater variability between participants (average Cronbach alpha = 0.600), in keeping with large individual differences in hedonic appraisals (Blood and Zatorre, 2001; Salimpoor *et al.*, 2009). Nevertheless, the high reliability of continuous judgments for arousal and global ratings above was sufficient to allow a direct comparison of behavioral results obtained outside the scanner with the neuroimaging data obtained in the fMRI experiment group. Moreover, as shown in Figure 1, continuous ratings for arousal and valence did not follow a similar time course and thus appeared not to share a large amount of variance. Therefore, these evaluations were not orthogonalized for further analyses.

Musical features

The factor analysis on all nine acoustic features extracted with the MIR toolbox (see Supplementary Table S2) showed that two factors are sufficient to describe 79.2% of the variance among these musical features. The factor loadings pointed to three distinct families of features that can be grouped together along these two main dimensions. One subgroup contains more timbre-related features (zero-crossing, centroid, brightness, entropy), whereas another group subsumes event-related or energy-dominated features (rms, event density, dissonance and tempo). Pulse clarity seems to constitute a unique feature distinct from both subgroups above. Adding the emotional ratings into a similar analysis showed a virtually identical factorial segregation, with two similar dimensions sufficient to explain 72.2% of the variance. This second analysis further indicated that arousal ratings grouped together with the event/energy-related musical features, whereas valence appeared to constitute a separate category closer to pulse clarity (see Figure 2).

Additional multiple regression analyses with the musical features and the emotional dimensions showed that arousal and valence ratings could be predicted by the nine musical features (see Supplementary Table S3). For arousal, all features except centroid significantly contributed to the prediction, whereas only the three features rms, pulse clarity and entropy showed a significant relation to valence.

fMRI results

Global ISCs over the entire time-course of music pieces showed that activity in auditory regions (HG, STG, PPt, PPo), as well as distributed networks in several cortical and subcortical structures, correlated significantly across participants. Figure 3 shows the three global ISC maps calculated for each musical piece at a FDR corrected significance threshold. As can be seen, ISC maps were highly similar for all three pieces. Besides

Table 1. ANOVAs of behavioral ratings as a function of test context and stimulus type

(a)

Global ratings	Main effect piece	Main effect context	Interaction
Attention to the music	$F(1.7, 77.3) = 3.12$, $P = 0.06$	$F(1,46) = 3.12$, $P = 0.049^*$	$F(1.7, 77.3) = 0.33$, $P = 0.72$
Distraction by environment	$F(2.0, 91.1) = 0.31$, $P = 0.39$	$F(1,46) = 7.05$, $P = 0.01^*$	$F(2.0, 91.1) = 1.83$, $P = 0.17$
Emotional reactions to the music	$F(1.9, 86.6) = 10.70$, $P < 0.0001^{***}$	$F(1,46) = 0.02$, $P = 0.89$	$F(1.9, 86.6) = 0.34$, $P = 0.70$
Valence	$F(1.6, 73.4) = 25.63$, $P < 0.0001^{***}$	$F(1,46) = 5.66$, $P = 0.02^*$	$F(1.6, 73.4) = 0.44$, $P = 0.60$
Familiarity	$F(1.8, 84.0) = 11.51$, $P < 0.0001^{***}$	$F(1,46) = 2.66$, $P = 0.11$	$F(1.8, 84.0) = 0.71$, $P = 0.82$

(b)

GEMS ratings	Main effect category	Main effect context	Interaction
Mendelssohn	$F(5.6, 258.3) = 6.33$, $P < 0.0001^{***}$	$F(1,46) = 1.49$, $P = 0.23$	$F(5.6, 258.3) = 2.29$, $P = 0.04^*$
Prokofiev	$F(4.9, 227.2) = 13.03$, $P < 0.0001^{***}$	$F(1,46) = 2.39$, $P = 0.13$	$F(4.9, 227.2) = 0.27$, $P = 0.93$
Schubert	$F(5.7, 261.3) = 6.13$, $P < 0.0001^{***}$	$F(1,46) = 1.84$, $P = 0.18$	$F(5.7, 261.3) = 0.45$, $P = 0.88$

$^{***}P < 0.0001$, $^*P < 0.05$.

Table 2. Reliability of continuous evaluations (Cronbach alpha)

	Arousal	Valence
Mendelssohn	alpha = 0.969	alpha = 0.365
Prokofiev	alpha = 0.976	alpha = 0.680
Schubert	alpha = 0.957	alpha = 0.748

Cronbach standardized alpha is reported

auditory regions, peaks of synchrony across the listeners arose bilaterally in occipital visual areas, posterior parietal cortex, premotor areas, cerebellum, medial prefrontal and orbitofrontal cortex, insula, caudate and ventral striatum. Correlations in the amygdala were only found for the pieces by Mendelssohn and Schubert. These results demonstrate that BOLD activity in these brain regions was driven in a 'synchronized' manner across individuals when listening to the same music pieces.

For the dynamic ISC analysis with moving time windows, a FDR corrected significance threshold for correlation values from each window was defined for each piece separately. As non-parametric resampling methods were used, the significance threshold differs between pieces (at $P = 0.05$, ISC values (r) with FDR correction were considered as significant when $r = 0.18$ for Mendelssohn; $r = 0.16$ for Prokofiev; $r = 0.16$ for Schubert (Pajula *et al.*, 2012). For time windows in which the ISC survived this threshold, the averaged BOLD signal time course was extracted from each ROI and was then correlated with the time course of emotion ratings, as well as with acoustic indices computed for each music piece (second-order correlation). The retransformed correlation coefficients and the corresponding P values resulting from one-sample t-tests, which test the correlations in the significant time windows against zero, are reported for the emotion ratings (arousal and valence) in Supplementary Table S4, and for the musical features in the Supplementary Table S5. For this analysis, an arbitrary threshold of $P = 0.001$ was adopted that allowed retaining the most reliable correlations between BOLD signal from synchronized brain regions and either acoustic or emotional dimensions of music, while minimizing false positives due to multiple testing. These analyses revealed significant second-order correlations for the two emotional dimensions that were evaluated continuously (arousal and valence), as well as for all nine musical features that were determined by MIR analysis. These results are described in the next sections and summarized graphically in Figures 4–8.

Emotional experience

We first tested for second-order correlations of brain activity (in synchronized brain regions) with the degree of emotional arousal and valence during each music piece, as obtained from the continuous rating procedure. For arousal, significant positive correlations were found in several bilateral auditory areas along STG, including planum temporal and polar. But also more critically, several subcortical regions correlated positively with arousal, notably the left insula, left amygdala, left thalamus and right caudate nucleus, as well as the right frontal opercular regions (Figure 4).

For valence, mainly negative correlations were found. These involved only a few secondary auditory regions on the left side, together with negative correlations in bilateral frontal operculum, left ACC and left post-central gyrus. Moreover, negative correlations were also found for subcortical regions including the left amygdala and right caudate. The only positive correlation with valence was found in the right SMA (see Figure 4 and Supplementary Table S5).

Thus, more pleasant moments in the music were associated with transient reductions in key areas of the limbic system at both cortical (ACC) and subcortical (amygdala and caudate) levels; whereas conversely subjective increases in arousal were associated with increases in partly overlapping limbic brain areas.

Acoustic features

As shown by the factor analysis above, the nine different musical features were grouped into a few distinct clusters of co-occurrence when the entire time course of the pieces was taken into account. The correlation analysis of brain activity (in synchronized areas) for each of these acoustic features confirmed a similar grouping at the anatomical level but also revealed slight differences. Brain areas responding to these acoustic features during moments of inter-subject synchrony are described below.

For the features centroid, brightness and zero-crossing, which describe the noisiness and homogeneity of the sound, significant positive correlations during ISC peaks were found in shared brain regions including bilateral auditory areas, bilateral inferior frontal gyrus, as well as superior and middle left frontal gyrus, plus left anterior cingulate cortex (see Figure 5 and Supplementary Table S5). Only zero-crossing also correlated with the left superior frontal sulcus. Additional positive
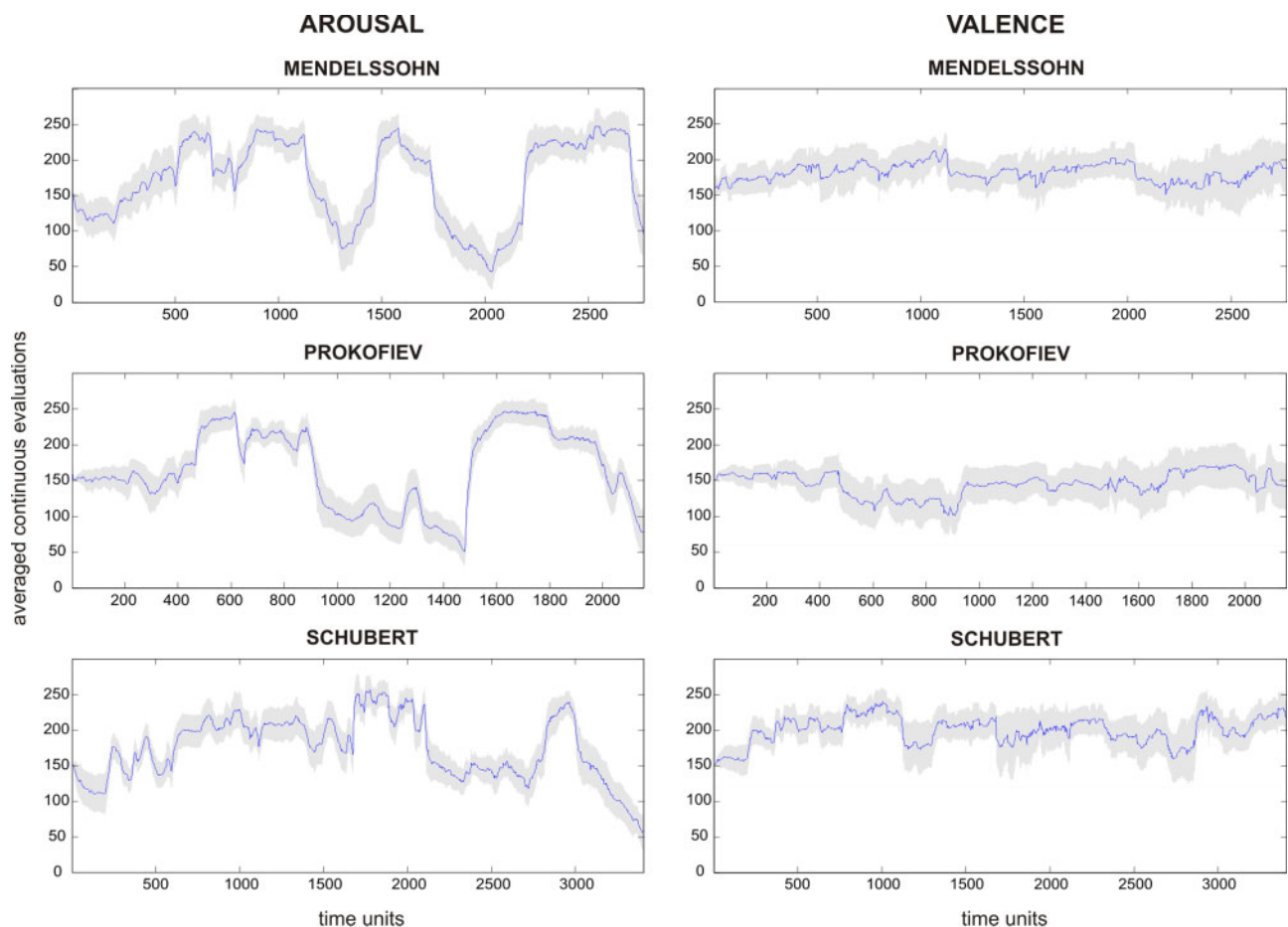
Fig. 1. Averaged evaluation scores over time for continuous emotional judgments of valence and arousal). The *x*-axis displays the time dimension sampled at 4 Hz and the *y*-axis shows the continuous evaluation scores (between 0 and 300). At the start position, the cursor was fixed to 150. These numbers were not visible to the participants. The gray-shaded areas represent 95% confidence intervals.

correlations were found for these features in subcortical regions including thalamus and bilateral putamen, while the left nucleus accumbens showed a negative correlation with the same features (see Figure 4 and Supplementary Table S4). Thus, activity in these regions tended to increase as a function of higher noisiness in music, except for the nucleus accumbens that showed a reduction instead.

A second group of features showing a common pattern of correlation during inter-subject synchrony was observed for entropy and event density. Besides auditory regions, these features correlated positively with activity in left pre-central gyrus, right IFG pars opercularis, left insula, thalamus (event density only right) and putamen (event density only left). In addition, for entropy, significant correlations were also found in left ACC and right precuneus (see Figure 6 and Supplementary Table S5), indicating that these regions were specifically sensitive to sound complexity. No negative correlation was observed.

Partly similar to the previous group, rms and dissonance appeared however to constitute a different subset of features, which we describe as energy related. Besides peaks in auditory and opercular frontal regions, we found positive correlations in bilateral thalamus, right precuneus and left insula. Furthermore, rms also showed selective positive correlations in the right putamen and right IFG pars opercularis, whereas dissonance was uniquely related to the right SMA. Conversely, negative correlations with these features were found in the left amygdala and left nucleus accumbens (see Figure 7 and

Supplementary Table S5). Thus, the louder and the more dissonant the music, the more activity in these limbic subcortical structures decreased when they synchronized between participants, but the more the activity increased in auditory areas, premotor circuits and insula.

Finally, the remaining two features, tempo and pulse clarity, showed slightly different correlation patterns as compared to other features (see Figure 8 and Supplementary Table S5). Notably, less correlation was observed in primary auditory areas. Instead, tempo significantly correlated with activity in bilateral IFG and right OFC. In subcortical regions, significant correlations occurred in left thalamus (for tempo only) and in putamen (for pulse clarity only left side), but none related to the limbic system.

Discussion

In this study, we investigated music perception and in particular musical emotion processing using ISC of brain activity, and then related the latter to the time course of different musical features. We observed significant synchrony across listeners in distributed brain networks including not only auditory cortices (as expected given exposure to similar auditory inputs) but also, more remarkably, several areas implicated in visual, motor, attention, or affective processes (see Figure 3). In particular, inter-subject synchrony occurred in key regions associated with reward and subjective feeling states such as the ventral
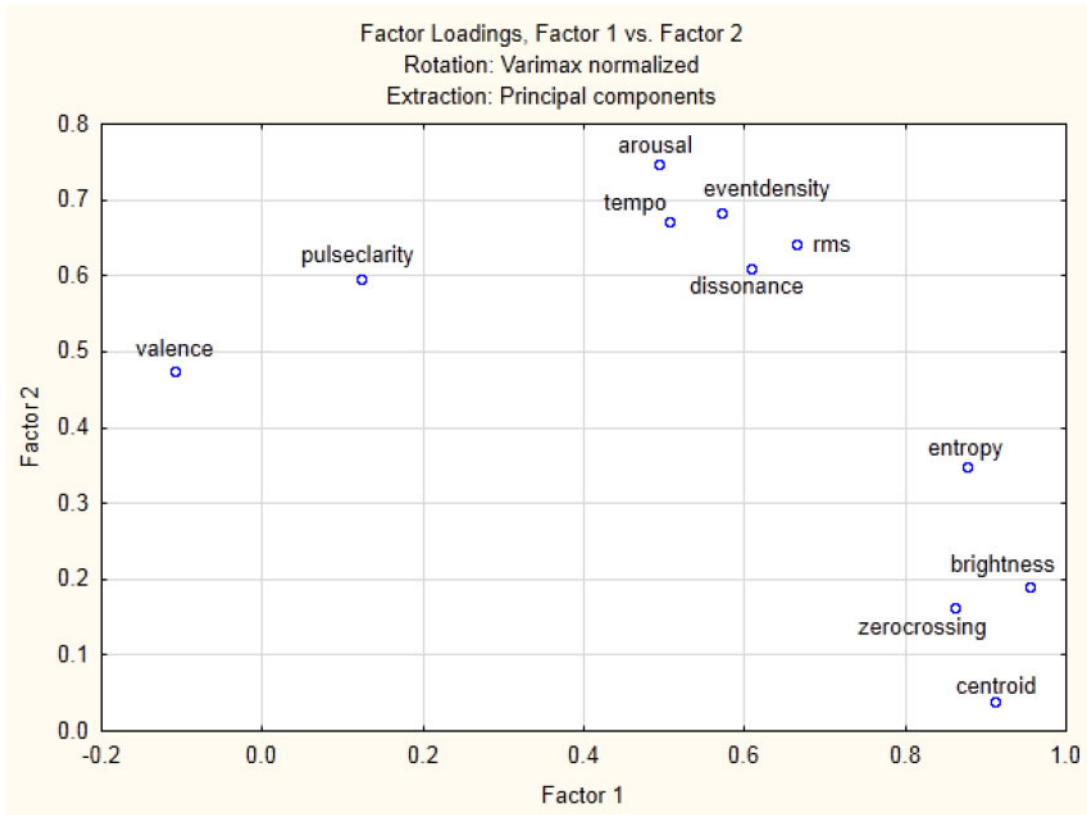
Fig. 2. Factorial analysis for all musical features combined with emotion ratings.
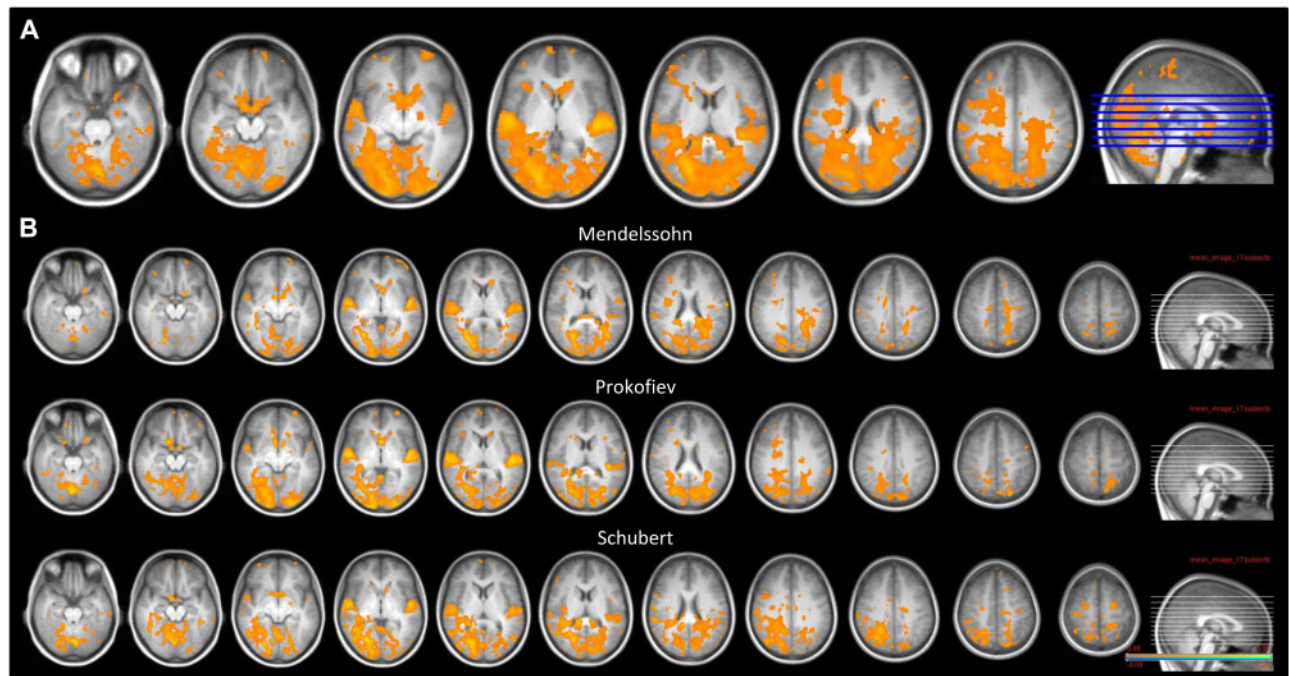

Fig. 3. (A) Global ISC maps overlaid on the mean structural brain image, combined for all three music pieces together. Significance threshold: $P < 0.05$ (FDR). (B) ISC maps for each music piece separately. Significance threshold: $P < 0.05$ (FDR), corresponding to a minimum threshold of $r = 0.08$ and a maximum of $r = 0.36$. In all cases, synchrony involved not only large areas in auditory cortex, but also visual areas, superior parietal areas, prefrontal areas, premotor cortex, cerebellum and ventral striatum.
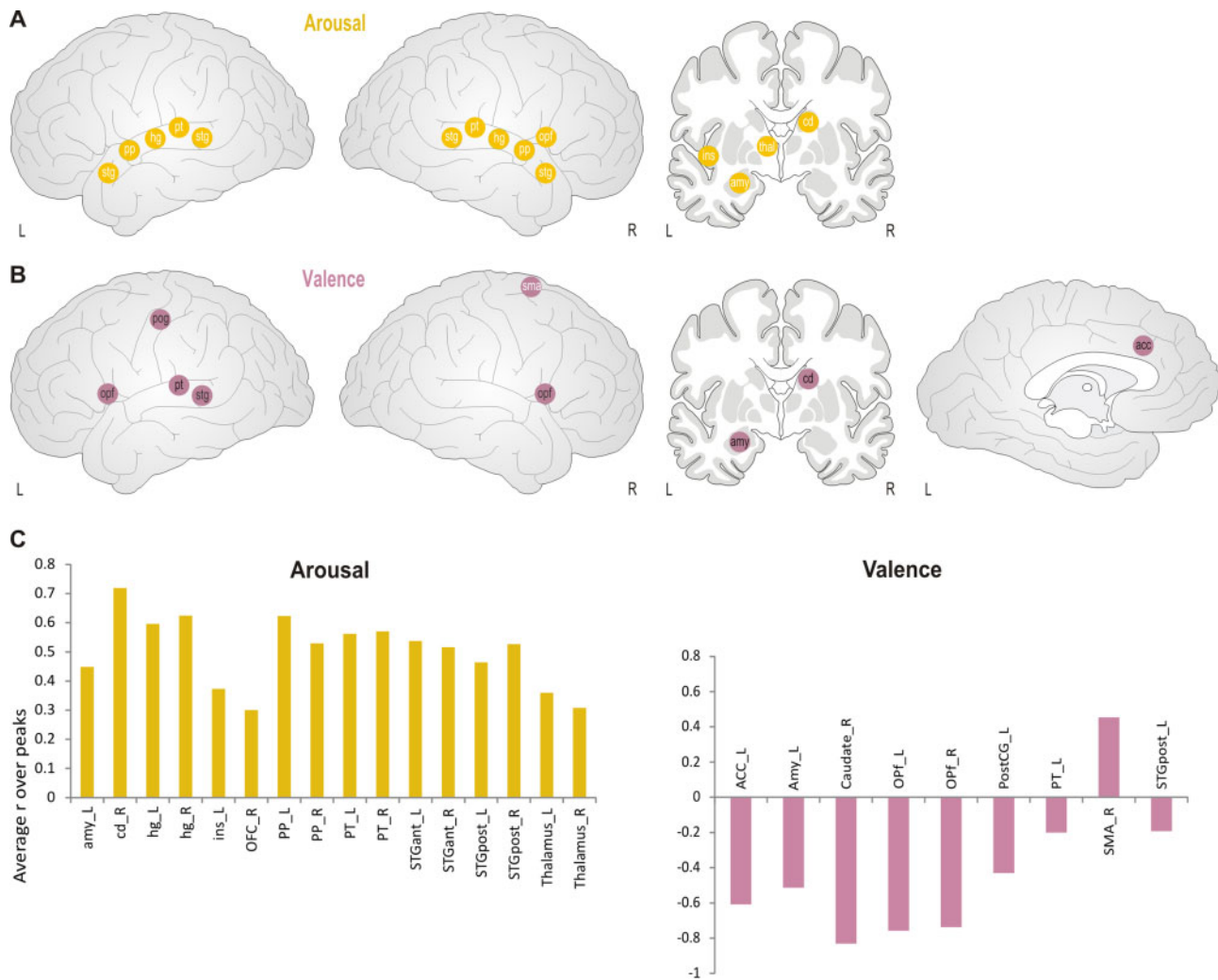
Fig. 4. Schematic correlation map for subjective emotional ratings, arousal (A) and valence (B). White letters indicate positive correlations, black letters indicate negative correlations. Abbreviations: acc: anterior cingulate, hg: heschl's gyrus, ins: insula, nac: nucleus accumbens, opf: frontal operculum, pog: postcentral gyrus, pp: planum polare, pt: planum temporale, sma: supplementary motor area, stg: superior temporal gyrus (anterior and posterior, respectively), thal: thalamus. (C) Averaged correlation coefficients are displayed for the regions which exceeded the significance threshold of $P = 0.001$.

striatum and insula. Synchrony in the amygdala was only found for the pieces of Mendelssohn and Schubert, which were the pieces which were more liked and induced stronger emotional reactions (Supplementary Figure S1). In addition, we found that brain activation during synchrony periods correlated with the presence of specific acoustic features in the music, with different features being associated with increases or decreases in different regions. Again, these effects involved widespread areas beyond auditory cortex, including cortical and subcortical limbic structures such as the insula and amygdala, the basal ganglia (caudate and putamen), prefrontal areas, as well as motor and premotor cortices. Moreover, these different musical features and the corresponding brain responses appeared to constitute different clusters, based on their time course (across the different music pieces) as well as according to their relation with subjective emotional judgments of arousal and pleasantness. The three main groups of musical features were timbre-related (centroid, brightness, zero-crossing), energy- or event- related (entropy, event density, rms, dissonance) and rhythm-related attributes (tempo, pulse clarity). Arousal effects showed activation patterns resembling mostly the energy- and rhythm-related features, whereas pleasantness shared activations with

both arousal and timbre features. These results accord with the fact that emotional reactions to music are presumably not elicited by single acoustical characteristics, but rather by a combination of different musical features.

Effects of emotional valence and arousal during ISC

The main goal of our study was to investigate whether inter-subject synchrony in brain activity elicited by music would involve regions associated with emotions. In addition, we sought to determine whether synchrony in specific brain regions would reflect particular acoustic features associated with musical emotions. Although several limbic regions and the basal ganglia were found to correlate with the subjective continuous ratings of arousal and valence, we did not observe a significant relation between activity of the nucleus accumbens (a region critically implicated in reward and pleasure; (Berridge and Robinson, 1998; Salimpoor *et al.*, 2013) and subjective reports of pleasantness. This contrasts with previous studies reporting activation of nucleus accumbens during peaks of pleasurable music (Blood and Zatorre, 2001, Menon and Levitin, 2005), but other studies on musical pleasantness have also
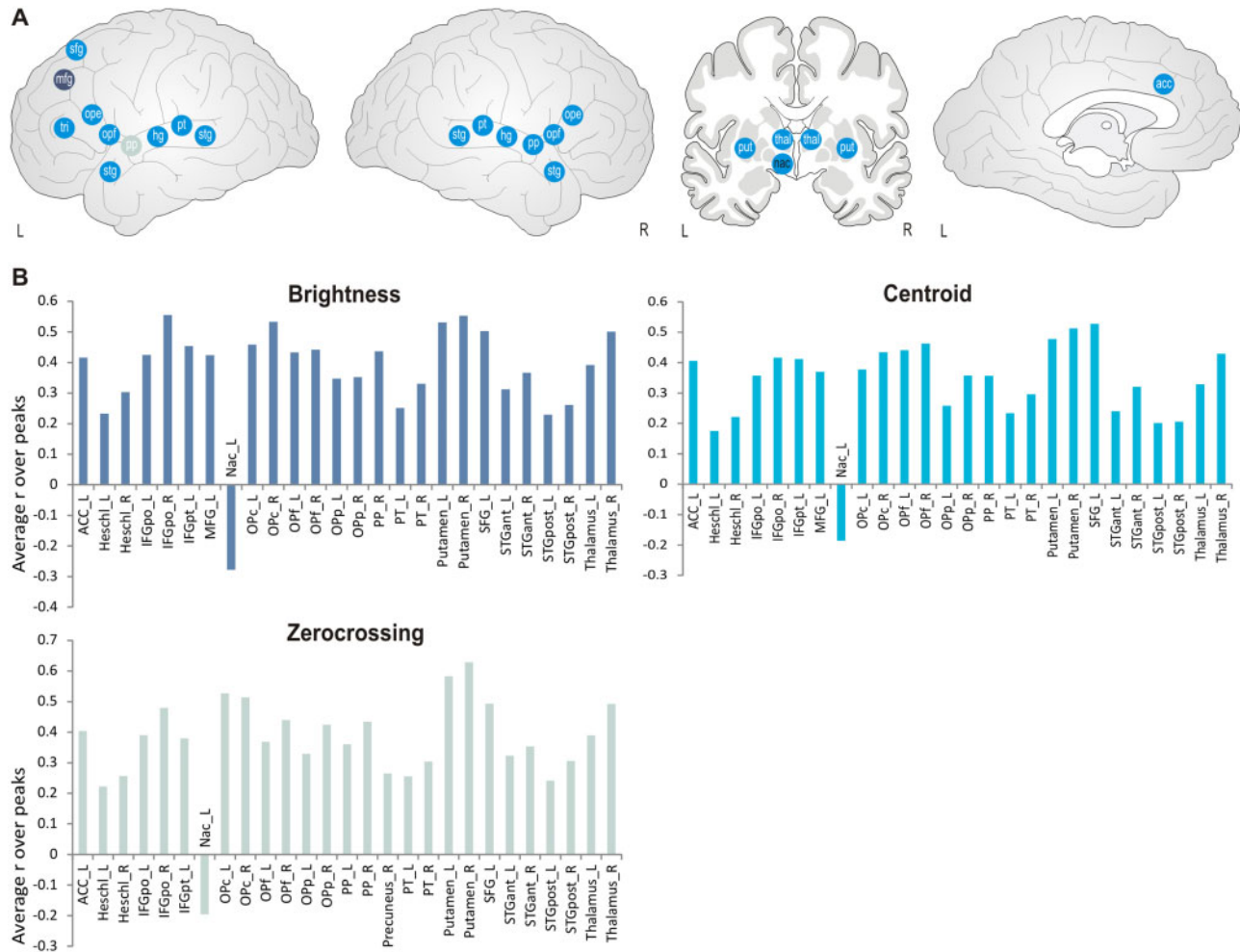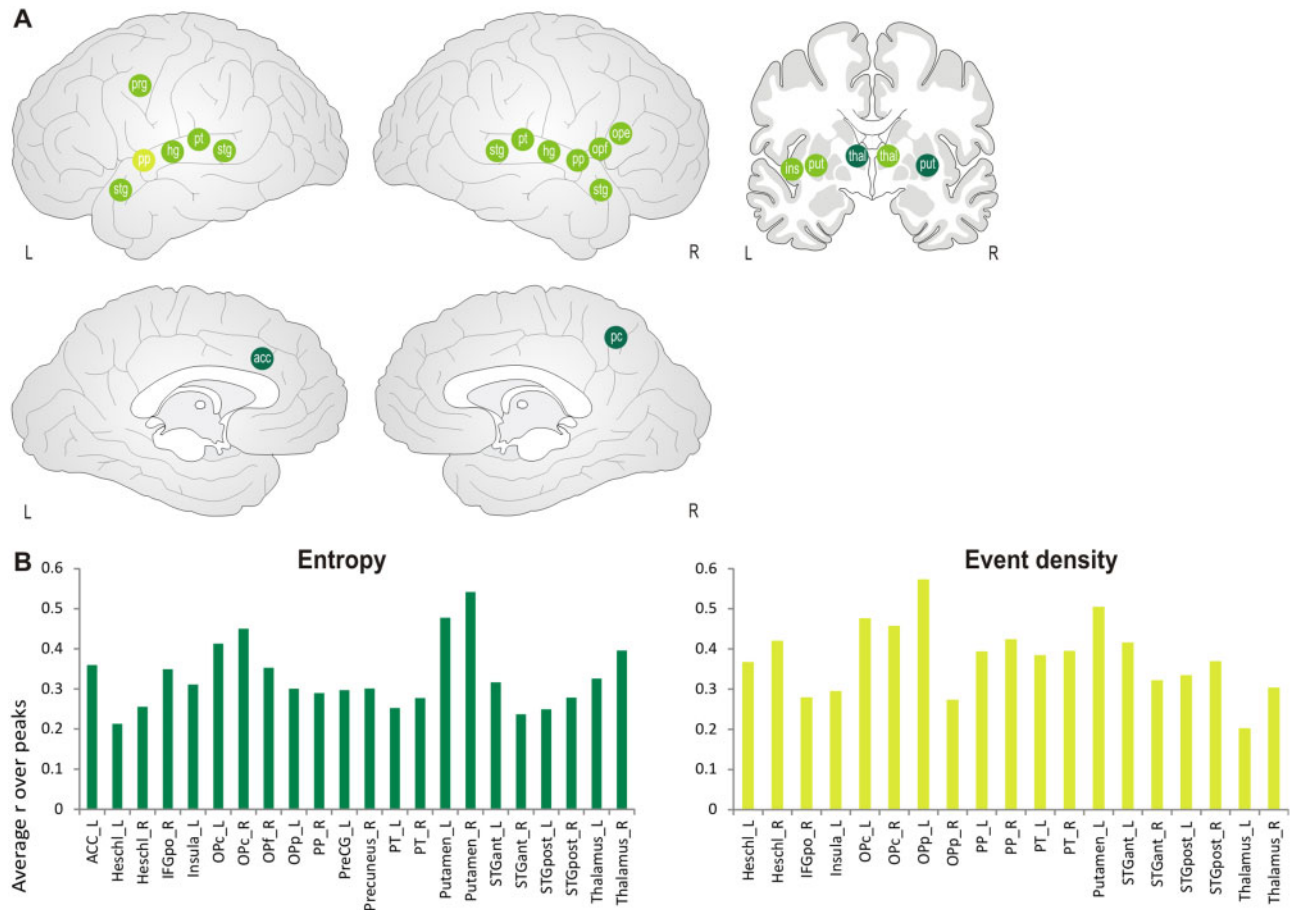
Fig. 5. (A) Schematic correlation maps for timbre-related features (centroid, brightness and zero crossing). Blue points indicate correlations common to all three features. Light blue points indicate correlation for zero-crossing only, dark blue points indicate correlations for all except for zero-crossing. White letters indicate positive correlations, black letters indicate negative correlations. Abbreviations: acc: anterior cingulate, hg: heschl's gyrus, mfg: middle frontal gyrus, nac: nucleus accumbens, ope: inferior frontal gyrus, pars opercularis, opf: frontal operculum, pp: planum polare, pt: planum temporale, put: putamen, sfg: superior frontal gyrus, stg: superior temporal gyrus (anterior and posterior, respectively), thal: thalamus, tri: inferior frontal gyrus, pars triangularis. (B) Averaged correlation coefficients are displayed for the regions which exceeded the significance threshold of $P = 0.001$.

failed to find consistent activation of the ventral striatum (Koelsch *et al.*, 2006). That we did not find positive correlations between activity of the nucleus accumbens and valence might be due to our experimental design, which primarily tested for coherent brain activity between participants over time. However, musical pleasantness represents a very subjective experience and it is possible that our participants did not always experience positive emotions at the exact same time but showed more idiosyncratic responses at various times, which would reduce the power of ISC measures. Nevertheless, as further discussed below, we found that nucleus accumbens correlated negatively with timbre- and energy-related features that reflect the acoustic frequency distribution, noisiness, and dissonance of music (see Figures 5 and 7).

In addition, the right caudate, another key component of the basal ganglia, showed a selective positive correlation with affective judgments of arousal but negative correlation with valence. Remarkably, unlike the accumbens, this structure was not modulated by any of the acoustic features, suggesting that it does not respond to low-level features but rather to higher level affective properties of the music. Previous studies reported that the caudate is more activated during happy than neutral

music (Mitterschiffthaler *et al.*, 2007), or found responses to both arousal and pleasantness (Colibazzi *et al.*, 2010). However, some studies also reported an involvement of the caudate for negative emotional experiences (Carretie *et al.*, 2009), or for the avoidance of positive emotions (Hare *et al.*, 2005). Here, we found the caudate showed opposite correlations for arousal (positive) and valence (negative). This finding converges with a recent study (Trost *et al.*, 2012) where we showed that the right caudate is not only sensitive to highly arousing pleasant emotions (e.g. power) induced by music, relative to low arousing pleasantness (e.g. tenderness), but also activates during highly arousing emotions with a negative valence that are characterized by subjective feelings of 'tension', that is, an emotion conveying a negative affect state of 'unease' (Zentner *et al.*, 2008; Trost *et al.*, 2012). Other studies indicate that the caudate is recruited during the anticipation of pleasant moments in the music (Salimpoor *et al.*, 2011) and mediates the automatic entrainment of motor or visuomotor processes by musical rhythms (Grahn *et al.*, 2008, Trost *et al.*, in press). Taken together, these data suggest that activity in the right caudate might jointly increase between participants in moments associated with the anticipation of specific musical events, such as those

Fig. 6. (A) Schematic correlation maps for event-related features (entropy and event density). Green points indicate correlations common to both features. Light green points indicate correlation for event density only and dark green points for entropy only. Abbreviations: acc: anterior cingulate, hg: heschl's gyrus, ins: insula, ope: inferior frontal gyrus, pars opercularis, opf: frontal operculum, pc: precuneus, pp: planum polare, prg: precentral gyrus, pt: planum temporale, put: putamen, stg: superior temporal gyrus (anterior and posterior, respectively), thal: thalamus. (B) Averaged correlation coefficients are displayed for the regions which exceeded the significance threshold of $P = 0.001$.

associated with a particular emphasis engaging attention of the listeners and corresponding to higher arousal, but inducing either positive or negative emotions depending on other concomitant parameters in the musical piece. This would accord with the role of the right caudate in encoding expectations or predictions of sensory or motor events in other domains (Balleine *et al.*, 2007; Asaad and Eskandar, 2011), as well as its implication in rhythmic perception and entrainment (Kokal *et al.*, 2011; Trost *et al.*, in press).

We found a similar relation between inter-subject synchrony and emotional effects in the left amygdala, with a positive correlation with arousal but negative correlation with valence judgments. This arousal effect accords with previous findings that the amygdala activates to moments in the music that are novel, alerting, and associated with structural changes (Fischer *et al.*, 2003; James *et al.*, 2008; Koelsch *et al.*, 2013), as well as to more intense stimuli in other sensory modalities (Anderson *et al.*, 2003; Winston *et al.*, 2005). The valence effect also converges with results showing that amygdala signals correlate negatively with the intensity of music-induced chills (Blood and Zatorre, 2001). Moreover, the amygdala shows greater response to unpleasant compared to pleasant music (Koelsch *et al.*, 2006) and is critically implicated in processing 'scary' music contents (Gosselin *et al.*, 2005; Gosselin *et al.*, 2007). Interestingly, our acoustic analysis demonstrated that synchronized activity in

the left amygdala was negatively related to the dissonance and rms features of music, which also correlated negatively with nucleus accumbens (see above), a pattern dovetailing with the study of Koelsch *et al.* (2006) who used dissonance to induce unpleasantness. Kumar *et al.* (2012) also suggested that the amygdala is differentially engaged by acoustical features associated with emotional unpleasantness. It is noteworthy that even though arousal is often closely linked with the loudness of the music, our results highlight a clear dissociation between these two factors, as we observed a positive correlation of amygdala activity with arousal but a negative correlation with rms. This finding indicates that the amygdala does not respond to the higher rms of music in a similar way as to loud events that may be arousing and behaviorally salient (Sander *et al.*, 2003). Thus, amygdala reactivity to rms and dissonance seems to reflect the emotional appraisal of the stimulus, rather than merely its acoustic energy.

The only positive correlation for valence ratings found in our study concerned the right SMA. As this region is a key component of motor initiation and sequence planning (Tankus *et al.*, 2009), our finding might indicate that musically induced pleasantness engenders motor action tendencies like the urge to dance or move with the music (Zatorre and Halpern, 2005). Moreover, induced mirth and smiling have been associated with activity in the SMA (Iwase *et al.*, 2002). Other studies of
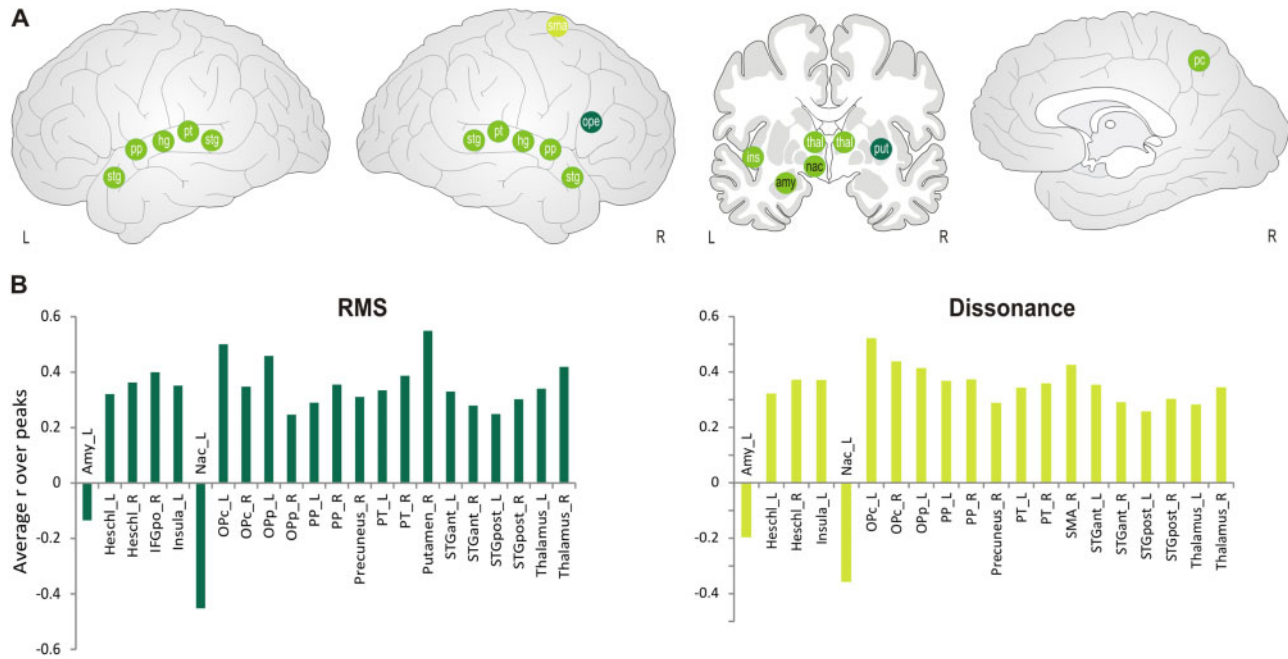
Fig. 7. (A) Schematic correlation map for energy-related features (rms and dissonance). Green points indicate correlations common to both features. Light green points indicate correlation for dissonance only, dark green points for rms only. White letters indicate positive correlations, black letters indicate negative correlations. Abbreviations: amy: amygdala, hg: heschl's gyrus, ins: insula, mfg: middle frontal gyrus, nac: nucleus accumbens, ope: inferior frontal gyrus, pars opercularis, opf: frontal operculum, pc: precuneus, pp: planum polare, pt: planum temporale, put: putamen, sfg: superior frontal gyrus, sma: supplementary motor area, stg: superior temporal gyrus (anterior and posterior, respectively), thal: thalamus, tri: inferior frontal gyrus, pars triangularis. (B) Averaged correlation coefficients are displayed for the regions which exceeded the significance threshold of $P = 0.001$.

music perception have observed a positive correlation between activation of SMA and the intensity of perceived chills (Blood and Zatorre, 2001). However, we also observed a selective correlation of right SMA with dissonance, which is rather associated with unpleasantness (Koelsch *et al.*, 2006). It is possible that motor activation in this case might reflect a covert reaction of avoidance (Sagaspe *et al.*, 2011), but further research is needed to clarify the significance of SMA increases during music processing and emotions.

Finally, apart from strong correlations with activity in sensory auditory areas and right opercular prefrontal regions, arousal also implicated the left insula and thalamus. These effects paralleled those of several acoustic features, particularly energy and event density-related features (i.e. dissonance, rms or entropy) that also strongly modulate insula activity. As these features correspond to more intense auditory signals, they are likely to contribute to enhance subjective arousal as reflected by such increases in sensory pathways and insula. These data accord with studies indicating that the insula play an important role in processing highly arousing and motivating stimuli (Berntson *et al.*, 2011). As our regression analysis of behavioral and acoustic data showed that the arousal values could be well predicted by the set of musical features (except from centroid), it is not surprising to find similar results also for the ISC analyses. One region also showing common correlations with different factors was the right opercular prefrontal cortex, which showed synchrony with increasing arousal as well as with all musical features except energy-related dimensions. This involvement of the frontal operculum might reflect a discriminative and analytic function of this region in auditory processing in general (Fruhholz and Grandjean, 2013). Interestingly, a negative correlation in this region was found for valence, which conversely might suggest a decrease of analytical hearing in moments of higher pleasantness.

## Effects of timbre-related features on brain activity and ISC

Besides uncovering intersubject synchrony in brain activity related to emotional experiences evoked by music, our study also identified neural patterns associated with the processing of specific music features. These features modulated several regions in auditory cortices as well as higher level cortical areas and subcortical areas.

For timbre-related features, which describe the spectral quality of auditory signals, we found differential activations in left inferior and superior frontal gyri, in addition to auditory cortex and subcortical regions. The left IFG including Broca's area has already been shown to be involved in timbre more than loudness processing for simple tones (Reiterer *et al.*, 2008). Activity in these lateral prefrontal areas might reflect abstract categorization and temporal segmentation processes associated with perception and aesthetic evaluation of the musical sound quality. Moreover, timbre-related features also correlated negatively with the left nucleus accumbens, a key part of the reward system influenced by dopamine and usually implicated in high pleasantness (Menon and Levitin, 2005; Salimpoor *et al.*, 2011). As timbre features measured here describe the noisiness and sharpness of sounds (see Supplementary Table S2), their impact on accumbens activity might reflect a relative unpleasantness of the corresponding musical events (i.e. indexed by elevated zero-crossing, high brightness and enlarged centroid).

Effects of event and energy-related features

For ISC driven by event- and energy-related musical features (entropy, event density, rms and dissonance), apart from auditory areas and thalamus, common increases were observed in the left insula for all four features and in the right precuneus for
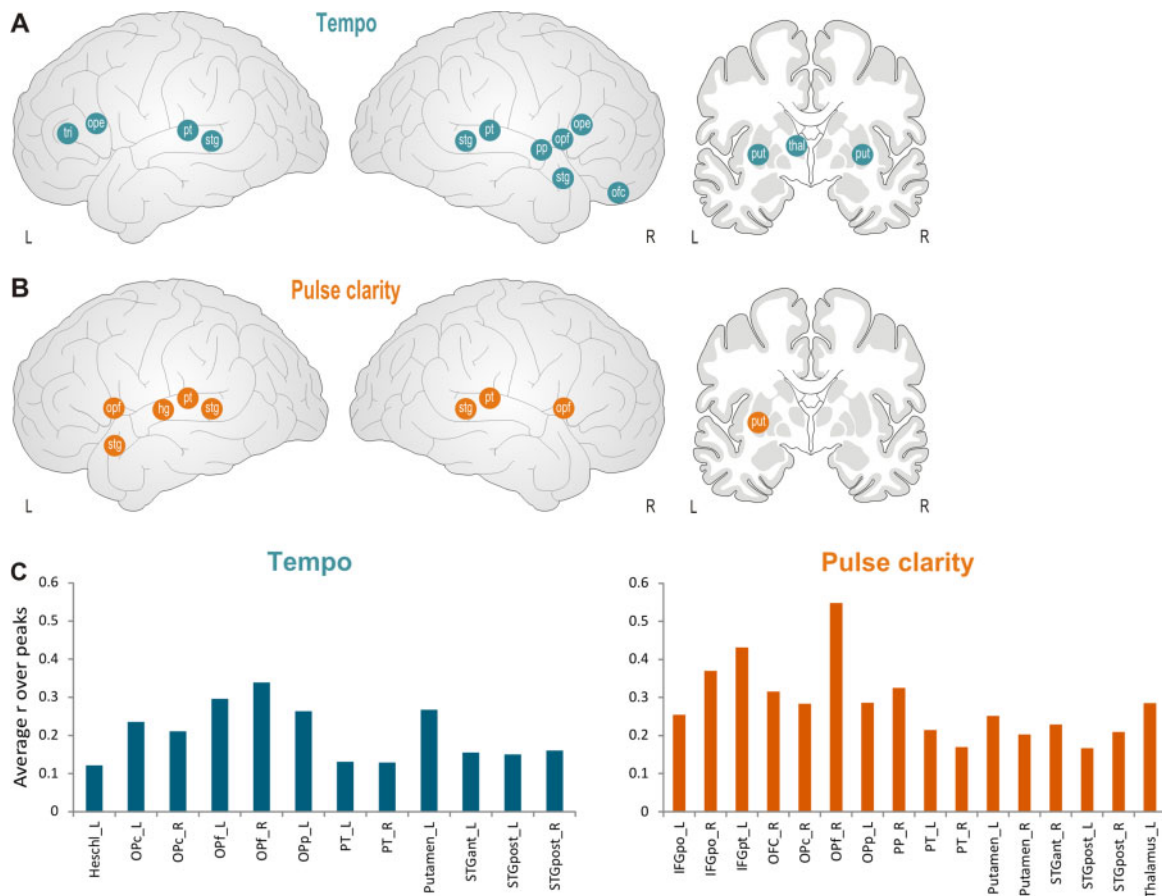
Fig. 8. Schematic correlation map for rhythm-related features, tempo (A) and pulse clarity (B). Abbreviations: hg: heschl's gyrus, ope: inferior frontal gyrus, pars opercularis, opf: frontal operculum, pp: planum polare, pt: planum temporale, put: putamen, stg: superior temporal gyrus (anterior and posterior, respectively), thal: thalamus, tri: inferior frontal gyrus, pars triangularis. (C) Averaged correlation coefficients are displayed for the regions which exceeded the significance threshold of $P = 0.001$.

all features except event density. While insula activity seemed related to arousal (see above), precuneus activity may correspond to heightened attentional focusing and absorption (Cavanna and Trimble, 2006). In support of the latter interpretation, music studies have reported that the precuneus is involved in the selective processing of pitch (Platel *et al.*, 1997), directing attention to a specific voice in harmonic music (Satoh *et al.*, 2001), and shifting attention from one modality to another (e.g. from audition to vision; (Trost *et al.*, in press). Moreover, in a previous study we found the precuneus to be selectively responsive to music inducing feelings of 'tension' (Trost *et al.*, 2012), an affective state that also heightens attention.

These four event and energy-related musical features could however be further distinguished. Entropy and event density correlated positively with the left precentral gyrus, which might encode action tendencies engendered by these two features, such as stronger motor tonicity associated with felt tension (Trost *et al.*, 2012), or instead stronger evocation of covert rhythmic movements due to increased event density in the music (Lahav *et al.*, 2007). On the other hand, rms and dissonance produced distinctive negative correlations with the left amygdala and left nucleus accumbens, reflecting a direct influence of these two features on emotion elicitation (Gabrielson and Juslin, 2003), as discussed above.

Effects of rhythm-related features

Both tempo and pulse clarity produced a distinct pattern, unlike other musical features. First, these features did not affect brain regions whose activity was associated with emotional valence or arousal ratings. Second, only a few correlations were found in auditory areas, none of which arose in the primary auditory cortex (except on the left side for pulse clarity). This suggests that these features constitute higher level properties, rather than low-level acoustic cues. The lack of effects in the right primary auditory cortex might also point to a lower temporal resolution relative to the left side (Zatorre and Belin, 2001). In addition, both, tempo and pulse clarity are features that do not change quickly, but rather vary over larger time scales or stepwise. In contrast, our ISC measures likely emphasized more transient activations. These factors might also account for the paucity of correlations with fMRI signal changes.

Notably, however, both tempo and pulse clarity modulated activity in the putamen, the portion of basal ganglia most closely linked to motor processes. Although the putamen was also modulated by other music features, it did not correlate with arousal and valence ratings. The basal ganglia play an important role in timing tasks and perception of temporal regularity (Buhusi and Meck, 2005, Geiser *et al.*, 2012). Moreover, the putamen is essentially involved in sensorimotor coordination

and beat regularity detection, whereas the caudate is more implicated in motor planning, anticipation, and cognitive control (Grahn *et al.*, 2008). In our study, only caudate activity correlated with emotion experience (see above), but not putamen.

Advantages and limitations of our method

The goal of this study was to identify synchronized brain processes between participants during music listening. We then also sought to determine which emotional dimensions or musical features could drive this synchronization. However, our approach did not aim at, and thus is not capable of, identifying more globally those brain structures that respond to particular emotional or acoustical features. Future studies correlating brain activity with the time-course of these features may provide additional information on the latter question (Chapin *et al.*, 2010; Alluri *et al.*, 2011; Nummenmaa *et al.*, 2012). Here, in addition to ISC, we used continuous judgments of valence and arousal, which represent a useful and sensitive tool to study emotional reactions over time (Schubert, 2010). Continuous judgments have been employed in other neuroimaging studies, allowing researchers to correlate the entire time-course of these judgments with brain activity (Chapin *et al.*, 2010; Alluri *et al.*, 2011; Abrams *et al.*, 2013). However, such analyses allow only the identification of brain regions that would track a specific aspect of the music to the same degree for the entire duration of the stimulus. This method is valid for constantly changing parameters such as low-level musical features. However, for emotional reactions, which instead occur episodically (but do not fluctuate continuously), these analyses seem less appropriate. Moreover, in standard GLM analyses, a characteristic linear response of the BOLD signal is assumed to reflect changes in stimulus features. However, in the case of music, meaningful events are more difficult to separate and may combine nonlinearly to produce ISC. Therefore our approach provides a useful tool to track emotional reactions and other neural processes responding to musical content independent of predefined effects on the BOLD signal.

Here, by correlating the fMRI signal with either emotion ratings or musical features in a restricted time window that corresponded to transient synchrony of brain activity across all participants, our measure of ISC should be sensitive to even brief and rare signal changes. On the other hand, this approach is less sensitive to slower variations and sustained states. Moreover, the size of the time window used for such analysis is critical. We chose a window of 10 scans (corresponding to 20 s in the music), which was a reliable compromise between sufficient statistical power to perform correlation analyses and adequate musical specificity without adopting a too long time period. This compromise was chosen to ensure a sufficient temporal resolution and to allow correlation analyses with a minimum of time points. However, more research is needed to validate this method more systematically, also with different window sizes. Further methodological work would also be required to compare activations identified by ISC and traditional GLM analysis based on the time-course of relevant features. It would also be interesting to explore ISC and temporal dynamics of more extended brain networks rather than single brain areas, and relate them to specific emotional responses. Future studies might also use ISC and reverse correlation techniques to identify more complex elements in music structures that predict synchronized responses and emotional experiences. In our study, an informal exploration revealed only partial correspondences between ISC peaks and theme transitions within each music pieces (see Supplementary Figure S3), but more advanced musicology measures might provide innovative ways to relate ISC in particular brain regions or networks to well-defined musical properties.

In any case, our results already demonstrate for the first time that ISC measures could successfully identify specific patterns of brain activity and synchrony among individuals, and then link these patterns with distinct emotion dimensions and particular music features.

Conclusion

By using a novel approach based on ISC during music listening, we were able to study the neural substrates and musical features underlying transient musical emotions without any *a priori* assumptions about their temporal onset or duration. Our results show that synchrony of brain activity among listeners reflects basic dimensions of emotions evoked by music such as arousal and valence, as well as particular acoustic parameters. ISC concomitant with emotional arousal was accompanied by increases in auditory areas and limbic areas including the amygdala, caudate and insula. On the other hand, ISC associated with higher pleasantness judgments was correlated with reduced activity in partly similar limbic areas, including the amygdala, caudate and anterior cingulate cortex, but with increased activity in premotor areas (SMA). Most of these effects were concomitant with variations in energy-related features of the music such as rms and dissonance. Further, our results reveal a dissociation in the left amygdala responses for the processing of rms and arousal (inducing decrease *vs* increased activity, respectively), which challenges the notion that sensory intensity and subjective arousal are always tightly linked (Anderson *et al.*, 2003). Rather, amygdala responses primarily reflect the affective significance of incoming stimuli which may be modulated by their intensity (Winston *et al.*, 2005). In addition, we found that the nucleus accumbens (a subcortical limbic region crucially implicated in reward and pleasure) showed no direct modulation by arousal or valence ratings, but a selective negative correlation with both energy and timbre-related features that describe the noisiness and spectral content of music. This pattern of activity in accumbens dovetails well with the fact that euphonic timbre and consonant harmony are typically associated with pleasurable music and positive emotions activating ventral striatum and dopaminergic pathways (Salimpoor *et al.*, 2011; Trost *et al.*, 2012).

Overall, these data point to a major role for several subcortical structures (in amygdala and basal ganglia) in shaping emotional responses induced by music. However, several cortical areas were also involved in processing music features not directly associated with musical emotions. Besides auditory areas recruited by a range of features, the left dorso-lateral and inferior prefrontal cortex was selectively sensitive to sound quality and timbre, whereas the insula and precuneus responded to both energy and event-related features that tend to enhance alertness and attention. Finally, rhythmical features differed from the others by their lack of correlation with the right primary auditory cortex, but predominant effects on inferior prefrontal areas and subcortical motor circuits (putamen). Taken together, these findings highlight that new methodological approaches using ISC can be usefully applied on various musical pieces and styles in order to better understand the role of specific brain regions in processing music information and their relation to specific emotional or cognitive responses.

Acknowledgements

Funding

Supplementary data

Supplementary data are available at *SCAN* online.

Conflict of interest. None declared.

References

Abrams, D.A., Ryali, S., Chen, T., *et al.* (2013). Inter-subject synchronization of brain responses during natural music listening. *European Journal of Neuroscience*, **37**, 1458–69.

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**, 3677–89.

Anderson, A.K., Christoff, K., Stappen, I., *et al.* (2003). Dissociated neural representations of intensity and valence in human olfaction. *Nature Neuroscience*, **6**, 196–202.

Asaad, W.F., Eskandar, E.N. (2011). Encoding of both positive and negative reward prediction errors by neurons of the primate lateral prefrontal cortex and caudate nucleus. *Journal of Neuroscience*, **31**, 17772–87.

Balleine, B.W., Delgado, M.R., Hikosaka, O. (2007). The role of the dorsal striatum in reward and decision-making. *Journal of Neuroscience*, **27**, 8161–5.

Berntson, G.G., Norman, G.J., Bechara, A., Bruss, J., Tranel, D., Cacioppo, J.T. (2011). The insula and evaluative processes. *Psychological Sciences*, **22**, 80–6.

Berridge, K.C., Robinson, T.E. (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Research Reviews*, **28**, 309–69.

Blood, A.J., Zatorre, R.J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 11818–23.

Blood, A.J., Zatorre, R.J., Bermudez, P., Evans, A.C. (1999). Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nature Neuroscience*, **2**, 382–7.

Buhusi, C.V, Meck, W.H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature Reviews Neuroscience*, **6**, 755–765.

Carretie, L., Rios, M., de la Gandara, B.S., *et al.* (2009). The striatum beyond reward: caudate responds intensely to unpleasant pictures. *Neuroscience*, **164**, 1615–22.

Chapin, H., Jantzen, K., Kelso, J.A., Steinberg, F., Large, E. (2010). Dynamic emotional and neural responses to music depend on performance expression and listener experience. *PLoS One*, **5**, e13812.

Cochrane, T. (2010). A simulation theory of musical expressivity. *The Australasian Journal of Philosophy*, **88**, 191–207.

Colibazzi, T., Posner, J., Wang, Z., *et al.* (2010). Neural systems subserving valence and arousal during the experience of induced emotions. *Emotion*, **10**, 377–89.

Etzel, J.A., Johnsen, E.L., Dickerson, J., Tranel, D., Adolphs, R. (2006). Cardiovascular and respiratory responses during musical mood induction. *International Journal of Psychophysiology*, **61**, 57–69.

Fischer, H., Wright, C.I., Whalen, P.J., McInerney, S.C., Shin, L.M., Rauch, S.L. (2003). Brain habituation during repeated exposure to fearful and neutral faces: A functional MRI study. *Brain Research Bulletin*, **59**, 387–92.

Fruhholz, S., Ceravolo, L., Grandjean, D. (2012). Specific brain networks during explicit and implicit decoding of emotional prosody. *Cerebral Cortex*, **22**, 1107–17.

Fruhholz, S., Grandjean, D. (2013). Processing of emotional vocalizations in bilateral inferior frontal cortex. *Neuroscience & Biobehavioral Reviews*, **37**, 2847–55.

Furman, O., Dorfman, N., Hasson, U., Davachi, L., Dudai, Y. (2007). They saw a movie: long-term memory for an extended audiovisual narrative. *Learning & Memory*, **14**, 457–67.

Gabrielson, A., Juslin, P.N. (2003). Emotional expression in music. In: Davidson, R. J. *et al.*, editors. *Handbook of Affective Sciences*, pp. 503–534. New York: Oxford University Press.

Geiser, E., Notter, M., Gabrieli, J.D. (2012). A corticostriatal neural system enhances auditory perception through temporal context processing. *The Journal of Neuroscience*, **32**, 6177–82.

Gomez, P., Danuser, B. (2007). Relationships between musical structure and psychophysiological measures of emotion. *Emotion*, **7**, 377–87.

Gosselin, N., Peretz, I., Johnsen, E., Adolphs, R. (2007). Amygdala damage impairs emotion recognition from music. *Neuropsychologia*, **45**, 236–44.

Gosselin, N., Peretz, I., Noulhiane, M., *et al.* (2005). Impaired recognition of scary music following unilateral temporal lobe excision. *Brain*, **128**, 628–40.

Grahn, J.A., Parkinson, J.A., Owen, A.M. (2008). The cognitive functions of the caudate nucleus. *Progress in Neurobiology*, **86**, 141–55.

Green, A.C., Baerentsen, K.B., Stodkilde-Jorgensen, H., Wallentin, M., Roepstorff, A., Vuust, P. (2008). Music in minor activates limbic structures: a relationship with dissonance? *Neuroreport*, **19**, 711–5.

Guhn, M., Hamm, A., Zentner, M. (2007). Physiological and musico-acoustic correlates of the chill response. *Music Perception*, **24**, 473–83.

Hare, T.A., Tottenham, N., Davidson, M.C., Glover, G.H., Casey, B.J. (2005). Contributions of amygdala and striatal activity in emotion regulation. *Biological Psychiatry*, **57**, 624–32.

Hasson, U., Furman, O., Clark, D., Dudai, Y., Davachi, L. (2008a). Enhanced intersubject correlations during movie viewing correlate with successful episodic encoding. *Neuron*, **57**, 452–62.

Hasson, U., Malach, R., Heeger, D.J. (2010). Reliability of cortical activity during natural stimulation. *Trends in Cognitive Sciences*, **14**, 40–8.

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science*, **303**, 1634–40.

Hasson, U., Yang, E., Vallines, I., Heeger, D.J., Rubin, N. (2008b). A hierarchy of temporal receptive windows in human cortex. *Journal of Neurosciences*, **28**, 2539–50.

Hevner, K. (1936). Experimental studies of the elements of expression in music. *American Journal of Psychology*, **48**, 246–68.

Iwase, M., Ouchi, Y., Okada, H., et al. (2002). Neural substrates of human facial expression of pleasant emotion induced by comic films: a PET Study. *NeuroImage*, **17**, 758–68.

Jaaskelainen, I.P., Koskentalo, K., Balk, M.H., et al. (2008). Inter-subject synchronization of prefrontal cortex hemodynamic activity during natural viewing. *Open Neuroimaging Journal*, **2**, 14–9.

James, C.E., Britz, J., Vuilleumier, P., Hauert, C.A., Michel, C.M. (2008). Early neuronal responses in right limbic structures mediate harmony incongruity processing in musical experts. *Neuroimage*, **42**, 1597–608.

Juslin, P.N. (2013). From everyday emotions to aesthetic emotions: towards a unified theory of musical emotions. *Physics of Life Reviews*, **10**, 235–66.

Juslin, P.N., Laukka, P. (2000). Improving emotional communication in music performance through cognitive feedback. *Musicae Scientiae*, **4**, 151–83.

Kanwisher, N., McDermott, J., Chun, M.M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience*, **17**, 4302–11.

Kauppi, J.P., Jaaskelainen, I.P., Sams, M., Tohka, J. (2010). Inter-subject correlation of brain hemodynamic responses during watching a movie: localization in space and frequency. *Frontiers in Neuroinformatics*, **4**, 5.

Kauppi, J.P., Pajula, J., Tohka, J. (2014). A versatile software package for inter-subject correlation based analyses of fMRI. *Frontiers in Neuroinformatics*, **8**, 2.

Koelsch, S. (2011). Toward a neural basis of music perception - a review and updated model. *Frontiers in Psychology*, **2**, 110.

Koelsch, S. (2013). Striking a chord in the brain: neurophysiological correlates of music-evoked positive emotions. In: Cochrane, T. et al., editors. *The Emotional Power of Music*, pp. 227–249. Oxford: Oxford University Press.

Koelsch, S. (2014). Brain correlates of music-evoked emotions. *Nature Reviews Neuroscience*, **15**, 170–80.

Koelsch, S., Fritz, T., Cramon, V., Muller, K., Friederici, A.D. (2006). Investigating emotion with music: an fMRI study. *Human Brain Mappings*, **27**, 239–50.

Koelsch, S., Rohrmeier, M., Torrecuso, R., Jentschke, S. (2013). Processing of hierarchical syntactic structure in music. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 15443–8.

Koelsch, S., Skouras, S. (2013). Functional centrality of amygdala, striatum and hypothalamus in a "Small-World" network underlying joy: an fMRI study with music. *Human Brain Mappings*, **35**, 3485–98.

Kokal, I., Engel, A., Kirschner, S., Keysers, C. (2011). Synchronized drumming enhances activity in the caudate and facilitates prosocial commitment–if the rhythm comes easily. *PLoS One*, **6**, e27272.

Krumhansl, C.L. (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimantal Psychology*, **51**, 336–53.

Kumar, S., von Kriegstein, K., Friston, K., Griffiths, T.D. (2012). Features versus feelings: dissociable representations of the acoustic features and valence of aversive sounds. *Journal of Neuroscience*, **32**, 14184–92.

Lahav, A., Saltzman, E., Schlaug, G. (2007). Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *Journal of Neuroscience*, **27**, 308–14.

Lartillot, O., Eerola, T., Toiviainen, P., Fornari, J. (2008). Multi-feature modeling of pulse clarity: Design, calidation, and optimization. In: Bello, J. P. et al. editors. *9th International Conference on Music Information Retrieval*, pp. 521–26. Philadelphia, PA.

Lartillot, O., Toiviainen, P. (2007) A Matlab toolbox for musical feature extraction from audio. In: *International Conference on Digital Audio Effects (DAFx-07)* Bordeaux, France.

Menon, V., Levitin, D.J. (2005). The rewards of music listening: response and physiological connectivity of the mesolimbic system. *Neuroimage*, **28**, 175–84.

Mitterschiffthaler, M.T., Fu, C.H., Dalton, J.A., Andrew, C.M., Williams, S.C. (2007). A functional MRI study of happy and sad affective states induced by classical music. *Human Brain Mappings*, **28**, 1150–162.

Nummenmaa, L., Glerean, E., Viinikainen, M., Jaaskelainen, I.P., Hari, R., Sams, M. (2012). Emotions promote social interaction by synchronizing brain activity across individuals. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 9599–604.

Pajula, J., Kauppi, J.P., Tohka, J. (2012). Inter-subject correlation in fMRI: method validation against stimulus-model based analysis. *PLoS One*, **7**, e41196.

Panksepp, J., Bernatzky, G. (2002). Emotional sounds and the brain: the neuro-affective foundations of musical appreciation. *Behavioural Processes*, **60**, 133–55.

Platel, H., Price, C., Baron, J.C., et al. (1997). The structural components of music perception - A functional anatomical study. *Brain*, **120**, 229–43.

Reiterer, S., Erb, M., Grodd, W., Wildgruber, D. (2008). Cerebral processing of timbre and loudness: fMRI evidence for a contribution of broca's area to basic auditory discrimination. *Brain Imaging and Behavior*, **2**, 1–10.

Sagaspe, P., Schwartz, S., Vuilleumier, P. (2011). Fear and stop: a role for the amygdala in motor inhibition by emotional signals. *Neuroimage*, **55**, 1825–35.

Salimpoor, V.N., Benovoy, M., Larcher, K., Dagher, A., Zatorre, R.J. (2011). Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nature Neuroscience*, **14**, 257–62.

Salimpoor, V.N., Benovoy, M., Longo, G., Cooperstock, J.R., Zatorre, R.J. (2009). The rewarding aspects of music listening are related to degree of emotional arousal. *PLoS One*, **4**, e7487.

Salimpoor, V.N., van den Bosch, I., Kovacevic, N., McIntosh, A.R., Dagher, A., Zatorre, R.J. (2013). Interactions between the nucleus accumbens and auditory cortices predict music reward value. *Science*, **340**, 216–19.

Sander, D., Grafman, J., Zalla, T. (2003). The human amygdala: An evolved system for relevance detection. *Reviews in the Neurosciences*, **14**, 303–16.

Satoh, M., Takeda, K., Nagata, K., Hatazawa, J., Kuzuhara, S. (2001). Activated brain regions in musicians during an ensemble: a PET study. *Brain Research Cognitive Brain Research*, **12**, 101–08.

Scherer, K.R. (2004). Which emotions can be induced by music? What are the underlying mechanisms? And how can we measure them? *Journal of New Music Research*, **33**, 239–51.

Scherer, K.R., Oshinsky, J.S. (1977). Cue utilization in emotion attribution from auditory stimuli. *Motivation and Emotion*, **1**, 331–46.

Schirmer, A., Kotz, S.A. (2006). Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences*, **10**, 24–30.

Schubert, E. (2004). Modeling perceived emotion with continuous musical features. *Music Perception*, **21**, 561–85.

Schubert, E. (2010). Continuous self-report methods. In: Juslin, P. N. and Sloboda, J. A., editors. *Handbook of Music and Emotion: Theory, Research, Applications*. Oxford: Oxford University Press.

Tankus, A., Yeshurun, Y., Flash, T., Fried, I. (2009). Encoding of speed and direction of movement in the human supplementary motor area. *Journal of Neurosurgery*, **110**, 1304–16.

Trost, W., Ethofer, T., Zentner, M., Vuilleumier, P. (2012). Mapping aesthetic musical emotions in the brain. *Cerebral Cortex*, **22**, 2769–83.

Trost, W., Fruhholz, S., Schon, D., *et al.* (in press). Getting the beat: entrainment of brain activity by musical rhythm and pleasantness. *Neuroimage*, **103**, 55–64.

Wilson, S.M., Molnar-Szakacs, I., Iacoboni, M. (2008). Beyond superior temporal cortex: intersubject correlations in narrative speech comprehension. *Cerebral Cortex*, **18**, 230–42.

Winston, J.S., Gottfried, J.A., Kilner, J.M., Dolan, R.J. (2005). Integrated neural representations of odor intensity and affective valence in human amygdala. *Journal of Neuroscience*, **25**, 8903–07.

Zatorre, R.J., Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex*, **11**, 946–53.

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

Zatorre, R.J., Halpern, A.R. (2005). Mental concerts: musical imagery and auditory cortex. *Neuron*, **47**, 9–12.

Zentner, M., Grandjean, D., Scherer, K.R. (2008). Emotions evoked by the sound of music: characterization, classification, and measurement. *Emotion*, **8**, 494–521.