

Cross-Cultural Music Phrase Processing: An fMRI Study

Yun Nan,^{1,2} Thomas R. Knösche,^{1,3*} Stefan Zysset,¹ and Angela D. Friederici¹

¹Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

²State Key Laboratory of Cognitive Neuroscience and Learning,
Beijing Normal University, Beijing, China

³Key Laboratory of Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing, China

Abstract: The current study used functional magnetic resonance imaging (fMRI) to investigate the neural basis of musical phrase boundary processing during the perception of music from native and non-native cultures. German musicians performed a cultural categorization task while listening to phrased Western (native) and Chinese (non-native) musical excerpts as well as modified versions of these, where the impression of phrasing has been reduced by removing the phrase boundary marking pause (henceforth called “unphrased”). Bilateral planum temporale was found to be associated with an increased difficulty of identifying phrase boundaries in unphrased Western melodies. A network involving frontal and parietal regions showed increased activation for the phrased condition with the orbital part of left inferior frontal gyrus presumably reflecting working memory aspects of the temporal integration between phrases, and the middle frontal gyrus and intraparietal sulcus probably reflecting attention processes. Areas more active in the culturally familiar, native (Western) condition included, in addition to the left planum temporale and right ventro-medial prefrontal cortex, mainly the bilateral motor regions. These latter results are interpreted in light of sensorimotor integration. Regions with increased signal for the unfamiliar, non-native music style (Chinese) included a right lateralized network of angular gyrus and the middle frontal gyrus, possibly reflecting higher demands on attention systems, and the right posterior insula suggesting higher loads on basic auditory processing. *Hum Brain Mapp* 29:312–328, 2008. © 2007 Wiley-Liss, Inc.

Key words: music perception; phrase structure; cross-cultural comparison; fMRI

INTRODUCTION

Auditory streams, be they speech or music, are structured into phrases, thereby facilitating their perception and further processing by the human brain. For the perception

of phrases, the correct detection of phrase boundaries is necessary. Such boundaries are indicated by specific cues. In the domain of music, these cues include small pauses, lengthening of the phrase final note, and the implicit tonal function of the phrase limiting tone [see e.g., Riemann, 1900]. While the most relevant of these cues is clearly the pause [Frankland and Cohen, 2004; Neuhaus et al., 2006; Riemann, 1900], the importance of other cues has been evidenced by several behavioral studies [see e.g., Boltz, 1993; Clarke, 1985; Jusczyk and Krumhansl, 1993]. Sensitivity to phrasing and phrase boundaries can already be observed in young infants, in the domains of both speech [Hirsh-Pasek et al., 1987] and music [Jusczyk and Krumhansl, 1993; Krumhansl and Jusczyk, 1990].

Recently, EEG and MEG correlates for the perception of phrase boundaries in music have been reported [Knösche

Contract grant sponsor: Deutsche Forschungsgemeinschaft; Contract grant number: KN 588/1-1

*Correspondence to: Thomas R. Knösche, Max-Planck-Institute for Human Cognitive and Brain Sciences, Stephanstrasse 1a, 04159 Leipzig, Germany. E-mail: knoesche@cbs.mpg.de

Received for publication 13 May 2006; Revision 9 February 2007; Accepted 18 February 2007

DOI: 10.1002/hbm.20390

Published online 11 May 2007 in Wiley InterScience (www.interscience.wiley.com).

© 2007 Wiley-Liss, Inc.

et al., 2005; Neuhaus et al., 2006]. In EEG, we identified a centro-parietal positive wave approximately 500 ms after the onset of the first tone following a phrase boundary. It bears some resemblance to the Closure Positive Shift (CPS), which indexes the processing of intonational phrase boundaries in speech [Pannekamp et al., 2005; Steinhauer et al., 1999]. It is therefore called *music CPS* and has been shown to be sensitive to a number of phrase boundary cues: pause length, length of the last tone preceding the pause, and harmonic function of this last tone [Neuhaus et al., 2006]. The same study demonstrated a relationship between the music CPS and the individual's formal musical training [Neuhaus et al., 2006].

Brain Networks Underlying Music Processing

In contrast to the behavioral and neurophysiological evidence mentioned above, there is very little knowledge on the neural networks underlying the processing of musical phrasing. Other aspects of music processing have received considerably more attention regarding their neuronal substrate by means of fMRI, MEG source localization, and lesion studies. In general, melody perception is reported to recruit the superior temporal [Kreutz et al., 2003; Morrison et al., 2003] and middle temporal gyri [Platel et al., 2003] as well as the frontal lobe [Bernal et al., 2004; Griffiths et al., 1998; Griffiths, 2003]. In addition, parietal, ventro-lateral prefrontal, and premotor areas [Halpern and Zatorre, 1999; Langheim et al., 2002; Zatorre et al., 1996] have also been evidenced to take a role in the process. The neural correlates for harmony processing seem to comprise a network including anterior cingulate cortex (ACC), temporal poles, parietal and occipital lobes, and the medial surface of the cerebellum bilaterally [Morrison et al., 2003; Satoh et al., 2001; Schmithorst, 2005]. The processing of musical syntax primarily involves left BA 44 (Broca's area) and its right hemisphere homologue, as shown by MEG [Maess et al., 2001] and fMRI [Koelsch et al., 2002]. fMRI research investigating temporal coherence reveals the importance of BA 47 for both linguistic [Dapretto and Bookheimer, 1999; Ni et al., 2000] and musical processing [Levitin and Menon, 2003, 2005].

Another much discussed aspect concerning the brain substrates of music processing is hemispherical asymmetry. Despite the suggestions from some brain-lesion and neuroimaging studies that music might engage the right hemisphere more than the left [Bernal et al., 2004; Halpern et al., 2004; Maess et al., 2001; Samson and Zatorre, 1994; Tramo and Bharucha, 1991; Zatorre et al., 1994], there are other studies that have shown a bilaterally distributed network for music processing [Kreutz et al., 2003; Meister et al., 2004; Platel et al., 2003; Tillmann et al., 2003].

To our knowledge, there is no evidence suggesting different brain loci for processing music of culturally familiar and unfamiliar styles, as suggested by the only published fMRI study on cross-cultural music perception [Morrison et al., 2003].

Objectives and Hypotheses in This Study

The goal of the present study is to specify: (1) the neural substrates of musical phrase boundary processing, (2) the possible differences in neural networks underlying the processing of culturally familiar and unfamiliar music in general, and (3) the interaction between 1 and 2, asking the questions of whether and how the processing of musical phrase boundaries might be influenced by the cultural familiarity of the music.

Brain networks of musical phrasing

Although there is, to our knowledge, only very little published brain imaging work concerned with the neural substrates underlying the processing of phrases in music or in speech prosody, ERP and MEG studies [Knösche et al., 2005; Steinhauer et al., 1999] suggest that there should be considerable activation of neuronal assemblies in response to phrase boundaries. Moreover, data from studies on related topics might allow the establishment of some hypotheses on the brain networks involved in the processing of phrase boundaries. These are studies that investigate the extraction of features from the acoustic input and those investigating integrative processing of sequential events.

A first possible candidate area for detecting phrase boundary cues and recognizing the presence of a phrase boundary is the planum temporale (PT). The PT contains the auditory association cortex, which is known to play a role in higher order processes of speech [Jäncke et al., 2002a] as well as processes that are common to speech and nonspeech input [Binder et al., 1996, 2000]. A recent study on the perception of normal speech, degraded speech accompanied with normal intonation, and flattened speech lacking the intonational contour [Meyer et al., 2004] led to the conclusion that the PT might support the integration of rapidly and slowly changing information. According to this view, the involvement of the PT in the processing of phrase boundaries in music as well as in speech would be very likely, since the integration of local phrase boundary markers with global information, like tonality and pitch contour, must be performed. Activation of the PT has also been observed in a number of studies on music perception associated with the recognition of musical features, like timbre, contour, key, interval, and rhythm [e.g. Liégeois-Chauvel et al., 1998; Menon et al., 2002; Ohnishi et al., 2001]. It could, therefore, be predicted that PT exhibits increased activity during the processing of phrase boundaries, supporting the integration of slow (global) information, which predicts the phrase boundary and rapid (local) information, which marks it. If, however, some of the phrase boundary markers are removed (e.g., the pause; see Methods section), there would be a conflict between both types of information and the identification of the phrase boundaries would require more effort. This could then be reflected by a further increase in PT activity. Indeed,

increased activity of the right PT associated with “flattened” speech (without prosody) as compared to normal speech has been reported recently [Friederici and Alter, 2004; Meyer et al., 2004].

The processing of integrational sequencing (timing and ordinality of events) involves structures of the so-called motor circuit, that is, medial and lateral premotor cortex, frontal operculum, striatum, and the intraparietal sulcus [Schubotz et al., 2000]. Within this network, the opercular part of the inferior frontal gyrus (IFG) has not only been associated with the processing of sequential rules in language syntax [Friederici, 2002], prosody [Meyer et al., 2004], and music syntax [Maess et al., 2001], but also with the perception of visually presented motion rhythms [Schubotz and von Cramon, 2001] and imagery of motion [Binkofski et al., 2000]. The orbital part of IFG has been linked to the processing of temporal structure in music [Levitin and Menon, 2003; Vuust et al., 2006]. More generally, the IFG seems to be involved in various processes, requiring the integration of information over time.

Hence, it appears that increased activity in various subregions of IFG as well as other constituents of the motor circuit has been observed, when the structure of sequences had to be processed. Since phrase boundary markers add to the amount of structure in a musical sequence, it would be reasonable to assume that the perception of phrased melodies causes increased activity in these regions.

In summary, we hypothesize that there will be increased activation of planum temporale, when the phrase boundary marking pause has been removed. Moreover, increased activity in constituents of the motor circuit and, in particular, of IFG is likely to be related to phrase boundary processing.

Cultural dependence of brain activation due to music processing

There has already been one attempt to isolate differences between processing native and non-native music [Morrison et al., 2003], which yielded a null result. There are two major methodological concerns with this study, which we strived to avoid in the current work. First, it has been acknowledged that there is great variability in individuals’ brain responses to music for people with different musicality levels [Aydin et al., 2005; Koelsch et al., 2005; Morrison et al., 2003; Patel et al., 1998; Satoh et al., 2001, 2003; Seung et al., 2005; Shahin et al., 2003], with or without the predisposition of absolute pitch [Barnea et al., 1994; Bermudez and Zatorre, 2005; Crummer et al., 1994; Itoh et al., 2005; Ohnishi et al., 2001; Peretz and Zatorre, 2005; Schlaug et al., 1995], and of different gender [Johnson et al., 1996; Koelsch et al., 2003a,b; Lee et al., 2003]. In order to specify the brain loci for certain functions by the fMRI methodology, a relatively strict selection of subjects is required in order to reduce intersubject variance. In the current study, only fairly young German female musicians without absolute pitch were recruited as participants. Secondly, rela-

tively long musical excerpts pose challenges to the restricted measuring time of an fMRI study. Very small trial numbers (e.g. 3 trials as in the study of Morrison et al. [2003]) might prevent the detection of effects [Huettel and McCarthy, 2001]. By using as many trials as possible (20 trials for each condition), the current study attempts to obtain a BOLD signal averaging output of good quality.

Although, as indicated above, there seems to be no published work reporting brain areas associated with the *cultural* familiarity of *music*, studies on the recognition of familiarity in general could help to generate hypotheses, if familiarity effects are domain general. Shah et al. [2001] identified the retrosplenial cortex as being associated with the processing of both familiar voices and faces in contrast to unfamiliar ones. The authors interpret their findings such that this part of the limbic system is involved in the recognition of familiarity. However, it is not clear whether the activation of the retrosplenial cortex relates to the detection of familiarity in general, or to the recognition of a particular voice/face, that is, to the identification of the particular person, or to emotional responses to the person personally known to the perceiver. Wheeler and Buckner [2004] as well as Iidaka et al. [2006] studied cortical networks associated with consciously remembering words presented about one day before. Iidaka et al. [2006] identified a bilateral network involving the middle frontal gyrus and inferior parietal lobule. The activity of the right parietal region in response to a repeated item was modulated by the repetition lag, leading the authors to the conclusion that this area would be critical for familiarity-based judgment [see also Yonelinas et al., 2005]. From these studies, we might hypothesize that right parietal cortex near the inferior parietal lobule and possibly the left parietal lobe and retrosplenial cortex might be also associated with familiarity mediated by cultural musical style.

An experiment targeting the familiarity of particular pieces of music showed activations in the left orbital inferior frontal gyrus and the superior temporal gyrus [Platel et al., 1997]. In a later study of the same group [Platel et al., 2003], semantic memory of melodies was associated with activation in medial and orbital frontal lobe, left anterior middle temporal gyrus (extending into the orbital part of the inferior frontal gyrus) as well as the left angular gyrus. It therefore seems that semantic memory of particular music pieces is associated with a mainly left hemispheric fronto-temporo-parietal network. This cognitive function, however, may differ from the more general familiarity based on a certain cultural style.

Since in musicians, familiar music is likely to be represented not only in the sensory, but also in the motor domain [Bangert and Altenmüller, 2003; Haueisen and Knösche, 2001], listening to culturally familiar (though not individually known) melodies might activate both motor and auditory brain areas. This might hold in particular, if the music belongs to the subjects’ own performance area (see, e.g., the comparison between pianists and chorus singers listening to piano music, studied by Haueisen and

Knösche [2001]). It would, therefore, be reasonable to expect greater activation in these areas for music pieces that are culturally familiar to the listener, all the more if the listeners are musicians and the culturally familiar music belongs to their own area of performance.

Apart from the cultural familiarity issue as such, we wanted to explore whether there are differences in the brain networks underlying phrase boundary perception as a function of cultural familiarity. From a previous ERP study [Nan et al., 2006], it seems that the processes of listening to culturally familiar and unfamiliar music are quite similar, with some modulation at early latencies, but not of the CPS. Hence, one goal of this study was to find out whether and to what extent this relative cultural universality also holds for the underlying neural networks.

The above formulated questions were addressed by presenting melodies of two different cultural styles (Chinese and Western), each in two versions (phrased/unphrased; see Methods section for details), to a group of German musicians.

METHODS

Participants

Twenty German female musicians with normal hearing abilities were paid for participation in this study. They did not have symptoms of any neurological, psychiatric, or internal disease, nor had they taken any medication for at least 3 days prior to the examinations. Because of their poor behavioral performance (hit rate lower than 80% for Western music), two subjects were excluded from the final analysis (as suggested by the repeated Grubb's outlier test at the level of $P < 1\%$). The remaining 18 participants were right-handed (average handedness: 97.7 ± 1.3) according to the Edinburgh Handedness Inventory [Oldfield, 1971]. None of them was an absolute pitch possessor according to the Absolute Pitch (AP) test (see below). The mean age was 23.7 ± 0.6 years (age range 19–28). The average starting age for instrument playing was 7.1 ± 0.5 years. Fourteen out of the eighteen participants reported that they were able to play at least two instruments. The present main instruments were the flute (7), piano (7), violin (3), and keyboard (1). Participants reported that they exercised 1.9 ± 0.3 h per day. All of the musicians had the opportunity to perform in public (on average, 7 times/year). They mainly played music from Baroque, Classical, and Romantic eras. This was the case for both training and spare-time playing. All musicians reported that they predominantly listened to classical music (10.2 ± 0.5 h/week). Informed written consent was obtained from participants prior to the investigation.

Absolute Pitch Test

The ability to identify the pitch of a given sound without any reference (absolute pitch) was assessed for all of the musicians with a mini-AP test introduced by Zatorre

[2003]. Each test sound was a computer-generated sine wave tone of 1 s duration with 50-ms attack and decay rates, and it was presented binaurally. All musicians got a training session consisting of 10–25 tones until they were familiar with the sound. The main AP test contained 10 tones, which were different from the training tones. They were presented twice in a randomized order with a 5-s interstimulus interval, within which the musicians were asked to write down the corresponding tone name. No reference tone was available to the musicians during the test. The response was regarded as correct if it was within a $1/2$ tone distance of the answer tone [Keenan et al., 2001].

Consistent with the prior-reported non-AP abilities, none of the musicians was able to perform the task. After the first few notes, 17 of them gave up the test; the other 3 musicians completed the task with low performance (correct response: 0, 0, 2 out of 10 testing items). This was in contrast to about 90% correct responses reported for AP musicians in literature [Keenan et al., 2001].

Materials

Forty short music pieces (20 pieces of Western music, 20 pieces of Chinese music) were chosen from pre-existing, although not well-known, melodies for the present study. These melodies (labeled as *phrased*) were clearly structured into two phrases (each with 4 bars), with pauses indicating phrase boundaries. The lengths of the excerpts ranged from 9 to 21 s. In order to obtain another set of matched *unphrased* melodies, the phrase boundary was modified by a professional musician by filling it with one or more musically plausible notes. This technique has already been used in earlier studies [Knösche et al., 2005; Nan et al., 2006; Neuhaus et al., 2006]. It ensures that the meter remains intact, which would not be the case after just eliminating the pause. It is clear that the manipulated pieces cannot be considered completely unphrased in the strict sense, since phrase boundary markers other than the pause (e.g., prefinal lengthening, harmonic closure, melodic contour, etc.) are still in place. Therefore, the condition might have been called “less phrased.” However, for the sake of brevity and consistence with earlier publications, we will call the experimental conditions *phrased* and *unphrased* hereafter. An additional set of melodies was created by combining two fragments from different examples, which, however, belonged to the same style (Western or Chinese). Fragments were created by splitting a musical piece into two at a randomly chosen splitting point. The *combined* pieces were always composed of a front and a tail fragment. These stimuli served as task items and were excluded from the final analysis. All of the melodies were synthesized with piano-like timbre.

Note that the change of timbre from a Chinese instrument (such as *erhu*) to a piano for Chinese music, though necessary, compromises Chinese music employed in the current study as entirely representative. As a result, Chinese music would not be recognized by its different

TABLE I. Comparison of phrase boundary features between Chinese and Western melodies

	Chinese melodies	Western melodies
Length of the preceding phrase (ms)	8,012 (441)	7,649 (349)
Pause length (ms)	549 (39)	500 (39)
Length of the whole melody (ms)	15,399 (850)	14,422 (640)

Values are averages, with 95% confidence intervals in parentheses.

timbre, but rather by its different tuning and scale system as compared with Western music. A brief summary of the acoustic features of both groups of stimuli is listed in Table I. Examples of stimuli are shown as scores in Figure 1.

Procedure

The experimental paradigm is sketched in Figure 2. All melodies were presented in pseudo-random order, in

order to avoid sequence effects. There was one experimental run for each participant. Each run consisted of 40 Western and 40 Chinese musical pieces (half *phrased* and half *unphrased*) together with 8 *combined* task items.

The task was to categorize the ongoing melodies into three classes: *combined* task melodies (which could be either Chinese or Western), Chinese melodies, and Western melodies. Before the real measurement, a detailed instruction and a training session were given to assure that the participants understood the task well. The answers were collected by delayed key-press (a visual prompt sentence was given after the presentation of each melody and also a feedback for the correctness of the answer was provided immediately). The same response pattern was assigned to each participant: right index key-press for Chinese, right middle for Western, and no key-press for the combined melodies. The combined items were introduced into the experiment to make the task more difficult and relying on the entire material and not only the beginnings. Since the combined items could be sewn together at any point, sub-

Chinese phrased (c38p)



Chinese unphrased (c38c)



Chinese phrased (c83p)



Chinese unphrased (c83c)



Western phrased (w46p)



Western unphrased (w46c)



Western phrased (w60p)



Western unphrased (w60c)



Combined



Figure 1.

Example scores of stimuli. The last line demonstrates how the combined target items are created. It is composed of a head fragment of w46p and a tail fragment of w60p. The seaming point is indicated by the interruption in the score.

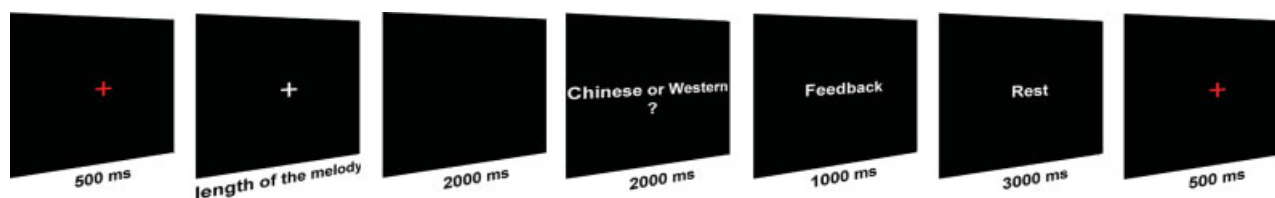


Figure 2.

Sketch of the experimental paradigm. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

jects had to attend to the end of a melody before making a decision. Subjects were not informed that the purpose of the experiment was the investigation of the perception of phrasing in a cross-cultural context.

MR Imaging

MRI data were collected by a 3T scanner (Siemens TRIO, Erlangen, Germany). Twenty axial slices (19.2 cm field of view (FOV) with a 64×64 matrix, 4 mm thickness, 1 mm interslice distance), parallel to the AC-PC plane and covering the whole brain were acquired using a single shot gradient recalled EPI sequence (TR 2000 ms, TE 30 ms, 90 flip angle). One functional run with 982 scans was carried out, with each scan sampling over the 20 slices. Before functional imaging, 20 anatomical T1-weighted MDEFT [Norris, 2000; Ugurbil et al., 1993] images (data matrix 256×256 , TR 1.3 s, TE 10 ms) and 20 T1-weighted EPI images with the same spatial orientation as the functional data were acquired. Additionally, in order to coregister functional scans, high-resolution whole-head 3-D MDEFT brain scans (128 sagittal slices, 1.5 mm thickness, FOV $25.0 \times 25.0 \times 19.2 \text{ cm}^3$, data matrix of 256×256 voxels) were acquired in a separate session.

Participants were comfortably placed in the scanner in supine position. Cushions were used to reduce head motion. The auditory stimuli were presented binaurally through specially constructed, MR compatible headphones (Commander XG, Resonance Technology, Northridge). The combination of soundproof headphones and special ear plugs attenuated the scanner noise without reducing the quality of sound stimulation. The loudness of the stimulation system was adapted so that the participants could perceive the stimuli without any problems (90 dB). The visual cues were displayed with an LCD projector on a back-projection screen mounted in the bore of the magnet behind the participant's head. Participants viewed the screen wearing mirror glasses.

Data Analysis

The software package LIPSIA was employed for MRI data analysis [Lohmann et al., 2001]. The series of processes included preprocessing, registration, normalization, and statistical evaluation of the MRI data. Preprocessing

was first started with an offline motion-correction of the functional data with the Siemens motion correction protocol (Siemens, Erlangen, Germany). Then a cubic-spline-interpolation and a temporal highpass filter with a cutoff frequency of 1/280 Hz were applied for an fMRI slice acquisition sequence and baseline corrections. The last step of preprocessing was spatially smoothing with a Gaussian kernel of 5.65 mm full-width half-maximum (FWHM). After the preprocessing, a rigid linear registration with six degrees of freedom (3 rotational, 3 translational) was performed in order to align the functional data slices onto a 3D stereotactic coordinate reference system. The rotational and translational parameters were acquired on the basis of the MDEFT [Ugurbil et al., 1993] and EPI-T1 slices to achieve an optimum match between these slices and the individual 3D reference data set. This 3D reference data set was acquired for each subject during a previous scanning session. The MDEFT volume data set with 160 slices and 1 mm slice thickness was standardized to the Talairach stereotactic space [Talairach and Tournoux, 1988], that is, scaled to match the extent of the Talairach brain and rotated and translated into the Talairach coordinate system with the origin at the anterior commissure (AC) and the anterior-posterior commissural line defining the y -axis. The rotational and translational parameters were then used to transform the functional slices using trilinear interpolation, so that the resulting functional slices were aligned with the stereotactic system.

Statistical evaluation was based on a least-squares estimation using the general linear model for serially auto-correlated observations [see also Aguirre et al., 1997; Friston et al., 1995; Worsley and Friston, 1995; Zarahn et al., 1997]. For each music style (Western, Chinese), a box-car function with the same duration as the music excerpts was convolved with a hemodynamic response function [HRF, constructed by a gamma density function; Glover, 1999] in order to generate the regressors. In order to model the process underlying the phrasing, an HRF with a response delay of 6 s, shifted by the latency of the phrase boundary offset, was used for the design matrix. Additionally, combined melodies were included in the model as an independent regressor (constructed the same way as for Western or Chinese melodies). The model equation, including the observation data, the design matrix and the error term, was convolved with a Gaussian kernel of dispersion of 4 s FWHM to account for

the temporal autocorrelation [Worsley and Friston, 1995]. For each participant, two contrast maps (i.e., estimates of the raw-score differences of the beta coefficients between specified conditions) were generated, which represented the main effects of: (1) Western versus Chinese music and (2) phrased versus unphrased melodies. As the individual functional datasets were all aligned to the same stereotactic reference space, the single-participant contrast-images were then entered into a second-level random effects analysis for each of the contrasts. The group analysis consisted of a one-sample *t*-test across the contrast images of all participants that indicated whether observed differences were significantly distinct from zero [Holmes and Friston, 1998]. Subsequently, *t* values were transformed to Z scores. Images were thresholded at $z > 3.09$ ($P < 0.001$, uncorrected). Moreover, a region was considered significant only if it contained a cluster of 5 or more continuous voxels [Braver and Bongiolatti, 2002; Forman et al., 1995]. This double threshold corresponds to a 5% multiple comparisons adjusted probability of falsely identifying one or more activated voxel clusters on the basis of Monte Carlo simulations (Alphasim/AFNI).

In addition, a time course analysis of the fMRI signal was conducted. Trial-averaged time courses (phrase boundary offset locked) for all of the four conditions (2 music styles \times 2 phrase conditions) were extracted from the preprocessed data for each participant at a sampling rate of 1 s. In order to detect the possible interaction between music style and phrasing conditions, a subsequent ROI analysis was conducted. All significant areas of both contrasts were employed as ROIs. The 4 s prior to phrase boundary onset was considered as baseline. Percent signal change was calculated relative to this baseline. Mean activation from the peak voxel over a 6 s time window (2 through 8 s after phrase boundary onsets) was computed for each participant. A two-way ANOVA with the factors STYLE (Western vs. Chinese) and COND (phrased vs. unphrased) was performed for each ROI.

RESULTS

Behavioral Results

We performed a two-way ANOVA with the factors STYLE (Western vs. Chinese) and COND (phrased vs. unphrased). All musicians performed better for the Western than for the Chinese melodies, as indicated by a significant main effect of STYLE ($F_{(1,17)} = 44.9$, $P < 0.001$). The mean recognition rate for Chinese melodies was 72.1% (standard error: 2.4%), while Western ones yielded a mean recognition rate of 90.6% (standard error: 1.1%). No significant main effect or interaction involving the phrasing condition (factor COND) was found.

Imaging Results

The locations of the clusters in the z-maps were identified in a Talairach scaled anatomical brain atlas [May et al., 2004].

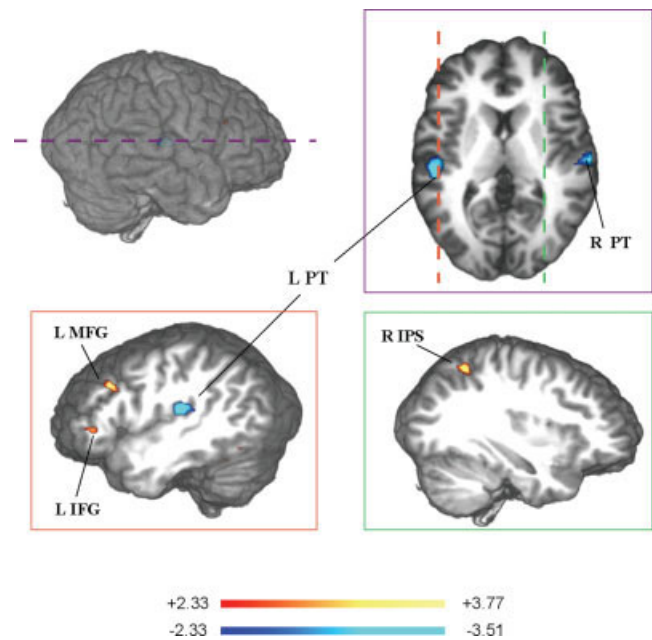


Figure 3.

Direct contrast maps of phrase vs. unphrased melody listening. Red colors indicate stronger activity for phrased, while blue colors indicate stronger activity for unphrased melodies. The abbreviations of the areas are the same as in Table II. In Figures 3 and 6, functional intersubject activation ($N = 18$) is superimposed onto a normalized 3D reference brain. All of the shown areas have their maximal Z-values above 3.09 ($P < 0.001$, uncorrected) and exceed a minimal size of 135 mm³.

Phrasing effect

Brain areas that exhibited significant differences between phrased and unphrased melodies are depicted in Figure 3 and listed in Table II. Listening to unphrased melodies resulted in increased activation in the bilateral supra-temporal plane, apparently comprising the anterior planum temporale (PT) and possibly involving the posterior Heschl's gyrus.¹ Time courses show that the effect is present in both Western and Chinese music for relatively anterior parts of the PT (Fig. 4A, second row), while somewhat more posteriorly and medially (Fig. 4B, second row) in the left hemisphere the phrasing effect shows up only for the Western music. The time courses also suggest an effect of musical style in that Western music causes higher activity, but this effect only proved significant for left posterior PT activation (see next subsection). On the other hand, phrased melodies produced greater levels of activation in

¹In the right hemisphere, the cluster was situated within PT, while in the left hemisphere, it was centered between Heschl's gyrus, PT, and parietal operculum.

TABLE II. Phrased vs. unphrased melodies

Location	Volume	Z-max	x	y	z
R IPS	189	3.95	28	-51	48
L MFG	162	3.82	-44	24	30
L IFG	162	4.01	-41	36	3
L PT	513	-3.6	-47	-21	15
R PT	135	-3.33	61	-18	9

R, right; L, left; IPS, intraparietal sulcus; MFG, middle frontal gyrus; IFG, inferior frontal gyrus, orbital part (BA 47); PT, planum temporale.

the right intraparietal sulcus, the left middle frontal gyrus, and an area in the orbital part of the left anterior IFG (BA 47, near the border to triangle part and to the anterior middle frontal gyrus). The time courses of the former two areas suggest a generally higher activation and also a greater contribution to the phrasing effect for the Chinese melodies. However, this was not significant.

Music style effect

The direct contrast between Chinese and Western music listening revealed the activations listed in Table III and depicted in Figure 5. Increased activation for Chinese music was found in the right lateralized network comprising the posterior insula as well as the middle frontal and angular gyri.² Processing of Western music was associated with stronger activation in a widely distributed network in both hemispheres. In the left hemisphere, two clusters of increased activation were found. One included motor areas, like the superior frontal gyrus (putative SMA) and the posterior precentral gyrus representing the putative mouth area of the primary motor cortex. The other cluster was located around the superior temporal gyrus and was extended into the PT and Heschl's gyrus. In the right hemisphere, increased activations were seen in the posterior bank of precentral gyrus in the classical hand area and the right homologue of the previously found mouth area. Moreover, an increased activation for Western music was observed in the right ventro-medial prefrontal cortex (VMPC).

Interactions between musical style and phrasing—ROI analysis

Among the observed regions found for the above two direct contrasts, significant interactions between COND (phrased/unphrased) and STYLE (Chinese/Western) were detected in three areas (see Fig. 6). For details, see the time

courses in Figure 4. An activity center near the left PT and posterior Heschl's gyrus (see Table III) showed a significant interaction between music style and phrase processing ($F_{(1,17)} = 4.40$, $P = 0.05$) in that differences between the phrased and unphrased conditions are present only for the Western musical style (according to post-hoc comparison, $P < 0.05$). Interactions were also observed along the left and right precentral gyri/central sulci in putative mouth and hand areas, respectively ($F_{(1,17)} = 4.45$, $P = 0.05$; $F_{(1,17)} = 7.84$, $P < 0.01$). In both areas, Western music seems to cause stronger activation when the phrase boundary was eliminated, while for Chinese music, the phrased versions were associated with stronger activation. Post-hoc comparisons, however, suggest that in the left hemisphere the interaction stems only from the Western melodies ($P < 0.05$ for Western, $P = 0.35$ for Chinese music), while in the right hemisphere there is no difference for Western melodies ($P = 0.53$), but a significant difference was observed for Chinese music ($P < 0.01$).

DISCUSSION

The present study aimed at (1) identifying the possible brain correlates for music phrase structure processing, and (2) isolating culture-dependent differences in the processing of music in general, as well as the processing of musical phrase boundaries in particular. With respect to these proposed research questions, the data revealed two lines of findings: (1) With respect to music phrase structure processing, bilateral planum temporale (PT) were identified as the most prominent brain areas for the processing of unphrased melodies as compared to phrased ones, while a fronto-parietal network consisting of left middle/inferior frontal gyri and right intraparietal sulcus was observed in response to phrased melodies. Interaction analysis revealed that part of left PT differences originate from the Western stimuli only. (2) Culturally unfamiliar, non-native (Chinese) music elicited more activation of a right-dominant network comprising the angular gyrus, posterior insula, and middle frontal gyrus, while familiar, native music style (Western music) activated a brain network comprising left superior temporal areas (PT, HeG, STG), right VMPC, and bilateral motor regions. An interaction between the cultural style of the music and phrasing conditions was revealed in some regions along the posterior bank of the precentral gyrus (significant in left putative mouth and right putative hand areas).

Phrase Structure Processing

The role of the planum temporale

There are two alternative interpretations of the bilaterally increased PT activation associated with unphrased melody processing. First, compared to the phrased melodies, the unphrased melodies contain a few more notes. Since the PT is known to be associated with basic acoustic

²Although these areas exceeded the significance threshold only in the right hemisphere, lowering the threshold of minimum cluster volume (135 mm²) also revealed activity in the homologous areas in the left hemisphere (left MFG consisted covered 27 mm², PI and AnG 81 mm² at the chosen z threshold of 3.09).

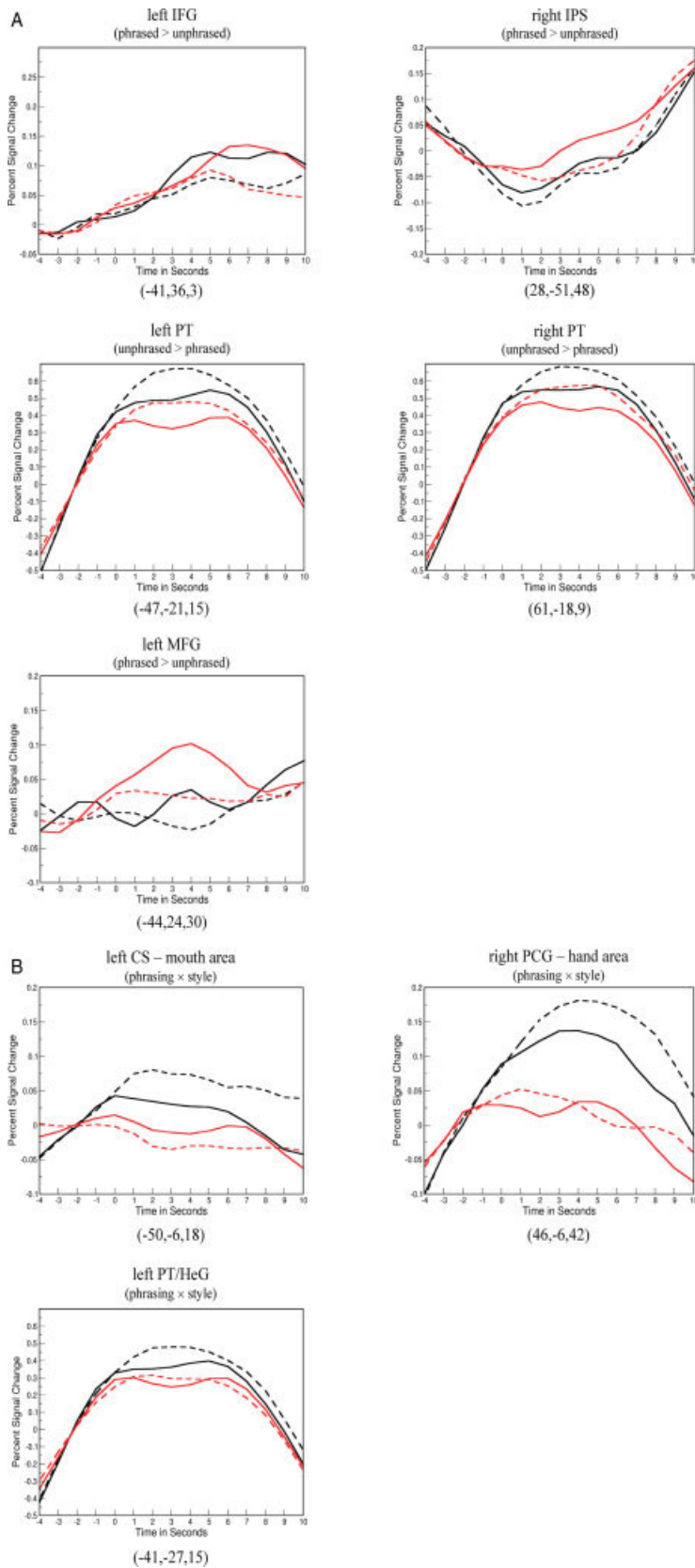


Figure 4.

Time courses for those areas with significant effects of the phrase boundary (contrast phrased–unphrased and interaction). Grey: Chinese; black: Western; solid: phrased; dotted: unphrased. The trigger point (time point zero) is at the offset of the phrase boundary marking pause. The bracketed number triples denote the Talairach coordinates of the activity centers. For abbreviations, see Tables II and III. **A**: Areas with significant contrast between phrased and unphrased items (see Table II). **B**: Areas with significant interaction between music style and phrasing processing from the ANOVA on the average values between 2 and 8 s after phrase boundary offset (see Fig. 6). [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

TABLE III. Western vs. Chinese music

Location	Volume	Z-max	x	y	z
R MFG ^a	432	-3.85	37	15	45
R AnG ^a	270	-3.61	43	-51	42
R PI ^a	189	-3.3	34	-21	3
L SMA/preSMA	594	3.73	-8	-9	54
L CS oral representation	2,511	4.62	-50	-6	18
R PCGhand representation	675	3.86	46	-6	42
R CS oral representation	567	3.9	55	-3	21
L PT/HeG	243	3.51	-41	-27	15
L PT/STG	297	3.78	-59	-24	15
R VMPC	135	3.58	10	42	0

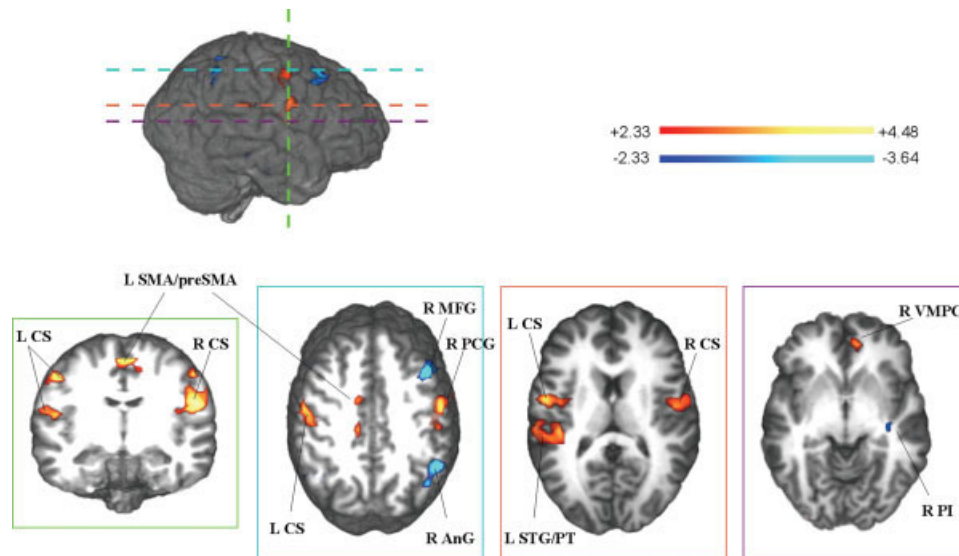
Tables II and III list Talairach coordinates [Talairach and Tournoux, 1988], Z-values, and volume (mm³) of the activated regions yielded from the averaged contrast images based on the individual contrast (Figure 5: Chinese vs. Western; Figure 3: phrased vs. unphrased). Z-values were thresholded at $|Z| > 3.09$ ($P < 0.001$, uncorrected). Only the activation clusters exceeding a minimal size of 135 mm³ (5 voxels) were listed.

^aThese areas are also active in the other hemisphere, but failed to reach the threshold for cluster size (for left PI and AnG, the blobs consisted of 3 voxels \equiv 81 mm², for left MFG 1 voxel \equiv 27 mm²). Abbreviations: R, right; L, left; MFG, middle frontal gyrus; AnG, angular gyrus/inferior parietal lobe; PI, posterior insula; PCG, medial part of precentral gyrus/foot area; CS, central sulcus, posterior bank of precentral gyrus in mouth and hand regions; PT, planum temporale; HeG, Heschl's gyrus; STG, superior temporal gyrus; VMPC, ventral medial prefrontal cortex (superior rostral gyrus).

processing [e.g. Griffiths and Warren, 2002; Knösche et al., 2003; Lütkenhöner and Steinsträter, 1998], this could have led to an increase in PT activity in our data.

Second, the involvement of bilateral PT in the processing of unphrased melodies might alternatively be explained by the recently proposed integration function between rapidly and slowly changing information [Meyer et al., 2004]. According to this view, the observed PT activity might reflect the identification of phrase boundaries triggered by local cues (rapid information) predicted on the basis of more global structural features of the music piece (e.g., contour, meter). The absence of a pause makes phrase boundary identification more difficult (but not impossible, because other cues are available, like e.g., prefinal lengthening), which results in stronger PT activity for the unphrased melody versions. If this interpretation is correct, one would predict that the observed effect is mainly or exclusively present with culturally familiar music. Indeed, at least for part of left PT, interaction analysis shows that increased activity for unphrased melody versions was only observed for the Western musical style.

More generally, the superior-posterior part of the bilateral temporal lobe, including the PT, has been suggested to constitute the primary substrate for constructing sound-based representations of speech [Hickok and Poeppel, 2000, 2004], but also for nonspeech sound processing [e.g. Jäncke et al., 2002b]. Auditory processing is generally asymmetric between hemispheres depending upon the specific material processed and the task. It is proposed that the left hemisphere specializes in temporal analysis and the right hemisphere specializes in spectral analysis [Zatorre 1997; Zatorre et al. 2002]. Similarly, Poeppel [2001, 2003] suggested the asymmetric sampling in time (AST) model, stating that the left hemisphere has a shorter time integration window (20–50 ms) and the right one has a

**Figure 5.**

Direct contrast maps of Western vs. Chinese melody listening. Red colors indicate stronger activity for Western, while blue colors indicate stronger activity for Chinese melodies. The abbreviations of the areas are the same as in Table III.

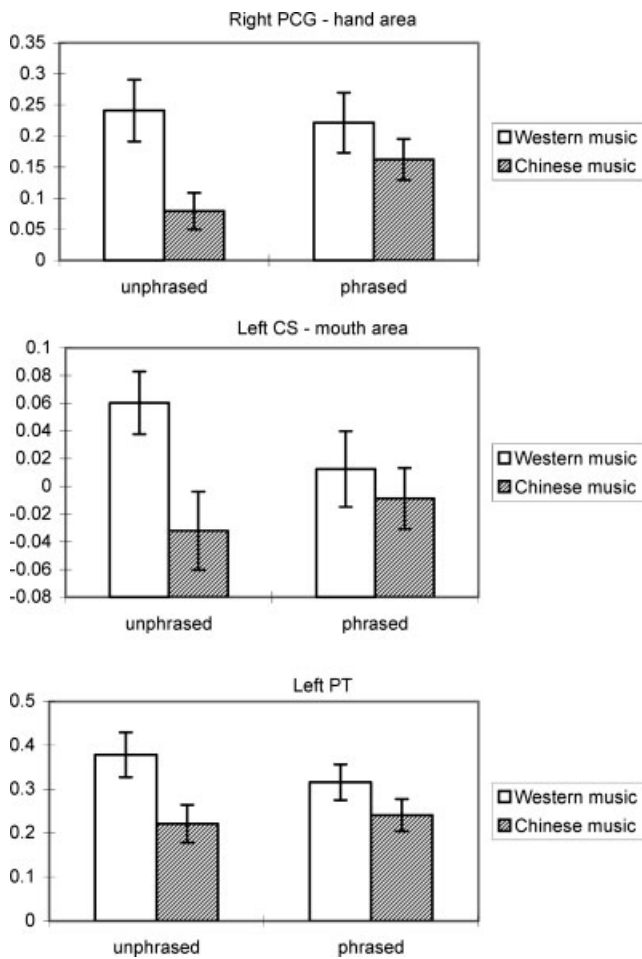


Figure 6.

Mean percent signal change and standard error for regions with significant interaction between music style and phrase processing. CS, central sulcus; PT, planum temporale.

longer time integration window (150–250 ms). In principle, phrasing and phrase boundary detection rely on both temporal and spectral cues. However, the distinction of very fine grained temporal features seems to play the most important role [see e.g. Neuhaus et al., 2006]. For example, the differences in length of the closing tones of the pre-boundary phrase was found to be about 10–70 ms on average, if the same musical pieces were played by a pianist with and without phrase boundary [Knösche et al., 2005; Table I]. Hence, according to the Zatorre model, phrasing should be processed preferentially in the left hemisphere. Indeed, in our data, the left PT seems to be considerably more active (see Table II), which is not only in accordance with the Zatorre model, but also with findings by Ohnishi et al. [2001], demonstrating left lateralized PT activity in musicians (but not in non-musicians) when listening to music.

A fronto-parietal network for phrase boundary processing

The processing of phrased, as contrasted to unphrased, melodies activated areas in the left orbital IFG (BA47), left middle frontal gyrus (MFG), and right intraparietal sulcus (IPS). The inferior frontal cortex has been shown to be involved in a wide range of brain processes including syntactic, semantic, and phonological aspects of speech processing [e.g. Caplan et al., 1999; Fiez et al., 1999; Poldrack et al., 1999; Wagner et al., 2001; see also Friederici and Alter, 2004; Friederici, 2002 for review], processing of nonlinguistic sequential input [Adams and Janata, 2002; Griffiths et al., 1999; Müller et al., 2001], music processing [Levitin and Menon, 2003, 2005; Maess et al., 2001], and processing of timing information [Schubotz and von Cramon, 2001; Schubotz et al., 2000]. The area observed in our study is well anterior to the bulk of IFG activations reported in literature. However, a similar focus of activation (distance 1.3 cm in Talairach space) has, e.g., been reported by Meyer et al. [2004] for the processing of degraded speech (all syntactic and semantic contents removed, while prosody is preserved) in contrast to normal and flattened (prosody removed) speech. Very similar activations have also been reported by Tillmann et al. [2003] in a musical priming experiment involving chord sequences (distances 1–2 cm to our activation). In the light of their own findings, as well as previous evidence, Tillmann et al. characterize the role of the inferior frontal cortices as "... processing and integrating information over time and the comparison of older stored information with newer incoming information." Such integration of information over time is also expected to be characteristic for the processing of phrase boundaries: the musical context established by the first phrase, comprising information on, for example, rhythm, meter, melodic line, timbre, and tonal key, has to be kept active over the time of the pause and then integrated with the new incoming information of the second phrase in order to establish a relation between phrases. It is worth pointing out that activity in BA47 has been found to specifically correlate with the processing of musical temporal structure [Levitin and Menon, 2003; Vuust et al., 2006], although there is some variability in the exact positions of the reported activity centers. The question remains, why is the activity in the left IFG *stronger* for the phrased as compared to the unphrased items. After all, one could argue that phrasing is more difficult without the pause and activity should be higher. However, the demand on working memory due to the temporal delay is higher in the phrased condition. It could therefore be that BA 47 reflects the working memory aspect of processing temporal structure in music. Indeed, it has recently been shown that the IFG (BA 45) plays a critical role in syntactic working memory during auditory language perception [Caplan et al., 1999].

The right IPS is part of the proposed dorsal fronto-parietal network of attention [Corbetta and Shulman, 2002;

Imaruoka et al., 2003]. This network is responsible for top-down signals for visual attention [Corbetta et al., 2002; Small et al., 2003; Thiel et al., 2004] and auditory attention [Collette et al., 2005]. In particular, Thiel et al. [2004] identified a network for reorienting attention in bilateral IPS and MFG, in which the loci of the left MFG and the right IPS activations are in close agreement with active areas found in our study (distance <1 cm in Talairach coordinates). MFG activation correlated to speech and tone segmentation processes has also been reported [Burton et al., 2000; LoCasto et al., 2004]. It is therefore likely that the observed fronto-parietal network (i.e. left MFG and right IPS) reflects more general cognitive processes like selective attention involved in the segmentation of the musical input into phrases.

Hence, the perception of phrased as compared to unphrased melodies in trained musicians is associated with higher activity in the orbital part of the left IFG and lower activity in bilateral PT. In principle, these effects are present for both culturally familiar and unfamiliar music, although some activity in the left PT is restricted to the familiar music condition. This pattern of culturally specific as well as culturally universal phrasing perception as indicated by these observed brain activations is in line with the findings of Nan et al. [2006] in the sense that music phrasing perception as reflected by the ERP also comprises culturally general (during the time range of the CPS) as well as cultural specific subcomponents (at earlier latencies).

Areas Associated With Music Style Processing

In general, it should be noted that there is a certain confound between the cultural familiarity of the music and basic acoustic properties, which are of course not (and cannot be) identical between the two musical styles. Therefore, we have to be cautious when interpreting activations in areas that are involved in basic acoustic processing, that is, the superior temporal lobe (Heschl's gyrus, PT and superior temporal gyrus).

Areas for familiar style

Among the areas found to be more active for the familiar music style (Western music), the role of the right ventral medial prefrontal cortex (VMPC) has been suggested to reflect the perceived ease of retrieval [Ranganath et al., 2003] or the "feeling of knowing" [Schnyer et al., 2004, 2005]. Consistent with these notions, VMPC showed enhanced activity for the familiar music style (Western music) as compared to the unfamiliar one (Chinese music) in the current study. Because in our case, the German musicians were not familiar with the musical pieces *per se*, the retrievable information would refer to the general Western musical schema rather than explicit knowledge of particular melodies. However, there is also evidence that

VMPC plays a role in pleasure and reward behaviors [Blood and Zatorre, 2001]. Hence, the observed VMPC response to Western music might alternatively be interpreted to indicate that the familiar style of music is more positively evaluated.

Some other constituents of the neural network for familiar music processing belong to the motor system, namely those in the bilateral precentral gyrus (including the SMA). The processing of event sequences in general (especially complex ones) is known to involve the so-called motor circuit [e.g. Schubotz et al., 2000; for a review see e.g. Janata and Grafton, 2003]. These processes are certainly influenced by attention and it seems plausible to assume that classical Western music, being more appealing to the German subjects in comparison to the Chinese examples, attracted more attention. Hence, the general attractiveness of culturally familiar music and consequently increased attention could be responsible for the observed activation increases in motor areas.

Alternatively, the observed motor activations corroborate the notion of "sensorimotor coupling." Music is not at all a purely perceptual phenomenon, but rather it is closely related to action, partly directed at producing the music (singing, instrument playing) and partly unrelated to production (clapping, tapping, swinging, dancing, etc.). Therefore, one possible explanation for the observed motor system activations might involve internal tapping or swinging, which is not directly related to the production of the perceived music. On the other hand, the relation between the perception of a stimulus and the motor behavior necessary to produce the same stimulus has been attributed to the mirror neuron system [Ferrari et al., 2003; Rizzolatti et al., 1996]. Activation in the human motor system is not only found during observation of action (classical mirror neurons), but also during listening of other's actions [audio-visual mirror neurons; Aziz-Zadeh et al., 2004], and during speech listening [echo neurons; Rizzolatti and Craighero, 2004; see also Fadiga et al., 2002; Sundara et al., 2001; Watkins and Paus, 2004; Watkins et al., 2003]. Thus, if there is an echo neuron system in humans, that is, a system that motorically resonates perceived verbal material [Rizzolatti and Craighero, 2004], it seems plausible to assume that other modes of human communication exhibit a similar phenomenon. Indeed, this has been demonstrated for humans viewing symbolic hand signs [Nakamura et al., 2004] and for pianists listening to music [Haueisen and Knösche, 2001].

In the present study, the Western musical pieces were not familiar to the listeners as such, but they were familiar in terms of cultural style. Hence, if the echo neuron account for the activity in the precentral gyrus holds, this means that cultural familiarity, that is, the familiarity with certain abstract rules and conventions, is sufficient to activate the echo neurons. Interestingly, Koelsch et al. [2006] observed activity in the Rolandic operculum near the putative representation of the vocal tract (larynx and pharynx) in non-musicians who were listening to emotionally pleas-

ant music. Hence, it might be that listening to music can trigger a perception-production integration system, which includes neural representations of vocal sound production in non-musicians and, in musicians, who are playing instruments, it also involves other parts of the motor system, namely hand representations.

Some of the areas of the PCG have been found to exhibit an interaction between phrasing and cultural familiarity. This indicates that explaining the observed differences between cultural styles only by internal playing, tapping or swinging is probably too simple.

In our study, most of the areas reported in literature as associated with familiarity based on semantic memory of pictures, faces, voices, or melodies (see Introduction) were not found to be activated. The reason for this may be that cultural familiarity of music does not rest on the contents of semantic or episodic memory, but on a system of abstract rules (i.e. implicit memory of a scheme). Both Western and Chinese melodies were unknown to the listeners. However, the Western pieces bear traits that make them sound familiar to Western subjects, like the harmonic (heptatonic) system, interval structure, and so forth.

Right hemisphere network for unfamiliar style

A right lateralized network was detected for unfamiliar music style (Chinese music), composed of areas in the middle frontal gyrus, angular gyrus (inferior parietal lobe), and posterior insula. The former two areas are known to reflect top-down aspects of selective attention [Corbetta and Shulman, 2002; Thiel et al., 2004]. It seems quite plausible that the demands on this cognitive faculty is different between the familiar Western and unfamiliar Chinese music. Although the right parietal cortex has also been found to be directly associated with familiarity based judgement [Iidaka et al., 2006; Yonelinas et al., 2005], those results would predict greater activity for the more familiar style, which would be in contrast to our findings.

We also found activation in the right posterior insula, a granular isocortical area associated with somatosensory, visual, and auditory functions [Mesulam and Mufson, 1985; Penfield and Faulk, 1955]. There is abundant evidence that this area is particularly involved in important aspects of auditory processing [for a review, see Bamiou et al., 2003]. First, research in monkeys indicates direct connections between the posterior insula, auditory thalamus, and the auditory cortex [e.g. Mesulam and Mufson, 1985]. Moreover, auditory responsive units were identified specifically in the posterior insula [Sudakov et al., 1971]. Second, human lesion studies demonstrate different forms of auditory agnosia and, in particular, amusia in patients with lesions involving the right posterior insular cortex [Fifer, 1993; Griffith et al., 1997; Habib et al., 1995; Spreen et al., 1965]. Third, there are a number of functional brain imaging studies suggesting the involvement of posterior insular cortex into basic auditory processing [e.g. Downar

et al., 2000; for a review see Bamiou et al., 2003]. We speculate that for the unfamiliar musical styles, such as basic auditory processes, for example, sound event detection and auditory Gestalt processing, are more demanding and result in increased activity in the right posterior insula.

CONCLUSION

The perception of phrased as compared to unphrased melodies in trained musicians is associated with higher activity in the orbital part of left IFG and lower activity in the bilateral PT. We suggest that the inferior frontal gyrus reflects working memory aspects necessary for the integration of information over time. The planum temporale might be involved in the identification of the phrase boundaries. In principle, these effects are present for both culturally familiar and unfamiliar music, although some activity in the left planum temporale is restricted to the familiar music condition.

When comparing the processing of culturally familiar to culturally unfamiliar music, we observe increased activity in parts of the motor system, demonstrating the possible existence of a human echo neuron system for music and strengthening the idea of sensori-motor coactivation during the perception of music, as put forward by, for example, Haueisen and Knösche [2001] and Koelsch et al. [2006]. Increased activity of the ventro-medial prefrontal cortex by culturally familiar music is in line with previous reports attributing a function regarding the “feeling of knowing” or “perceived ease of information retrieval” (i.e. familiarity) to this area [Ranganath et al., 2003; Schnyer et al., 2004, 2005].

Moreover, in both contrasts (phrased vs. unphrased and culturally familiar vs. unfamiliar), brain areas belonging to the fronto-parietal attention network were found, suggesting a greater involvement of more general cognitive functions, in particular selective attention, in the processing of phrase boundaries as well as unfamiliar musical sequences in contrast to familiar ones.

REFERENCES

- Adams RB, Janata P (2002): A comparison of neural circuits underlying auditory and visual object categorization. *Neuroimage* 16:361–377.
- Aguirre GK, Zarahn E, D'Esposito M (1997): Empirical analyses of BOLD fMRI statistics 2. Spatially smoothed data collected under null-hypothesis and experimental conditions. *Neuroimage* 5:199–212.
- Aydin K, Ciftci K, Terzibasoglu E, Ozkan M, Demirtas A, Sencer S, Minareci O (2005): Quantitative proton MR spectroscopic findings of cortical reorganization in the auditory cortex of musicians. *Am J Neuroradiol* 26:128–136.
- Aziz-Zadeh L, Iacoboni M, Zaidel E, Wilson S, Mazziotta J (2004): Left hemisphere motor facilitation in response to manual action sounds. *Eur J Neurosci* 19:2609–2612.

- Bamiou D-E, Musiek FE, Luxon LM (2003): The insula (Island of Reil) and its role in auditory processing—Literature review. *Brain Res Rev* 42:143–154.
- Bangert M, Altenmüller EO (2003): Mapping perception to action in piano practice: A longitudinal DC-EEG study. *BMC Neuroscience* 4:26.
- Barnea A, Granot R, Pratt H (1994): Absolute pitch—Electrophysiological evidence. *Int J Psychophysiol* 16:29–38.
- Bermudez P, Zatorre RJ (2005): Conditional associative memory for musical stimuli in nonmusicians: Implications for absolute pitch. *J Neurosci* 25:7718–7723.
- Bernal B, Altman NR, Medina LS (2004): Dissecting nonverbal auditory cortex asymmetry: An fMRI study. *Int J Neurosci* 114:661–680.
- Binder JR, Frost JA, Hammeke TA, Rao SM, Cox RW (1996): Function of the left planum temporale in auditory and linguistic processing. *Brain* 119:1239–1247.
- Binder JR, Frost JA, Hammeke TA, Bellgowan PSF, Springer JA, Kaufman JN, Possing ET (2000): Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex* 10:512–528.
- Binkofski F, Amunts K, Stephan KM, Posse S, Schormann T, Freund H-J, Zilles K, Seitz RJ (2000): Broca's region subserves imagery of motion: A combined cytoarchitectonic and fMRI study. *Hum Brain Mapp* 11:273–285.
- Blood AJ, Zatorre RJ (2001): Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proc Natl Acad Sci USA* 98:11818–11823.
- Boltz MG (1993): Time-estimation and expectancies. *Mem Cogn* 21:853–863.
- Braver TS, Bongiolatti SR (2002): The role of frontopolar cortex in subgoal processing during working memory. *Neuroimage* 15:523–536.
- Burton MW, Small SL, Blumstein SE (2000): The role of segmentation in phonological processing: An fMRI investigation. *J Cogn Neurosci* 12:679–690.
- Caplan D, Alpert N, Waters G (1999): PET studies of syntactic processing with auditory sentence presentation. *NeuroImage* 9:343–351.
- Clarke E (1985): The musical mind: The cognitive psychology of music, Sloboda JA. *Nature* 315:696. Book review.
- Collette F, Van der Linden M, Laureys S, Delfiore G, Degueldre C, Luxen A, Salmon E (2005): Exploring the unity and diversity of the neural substrates of executive functioning. *Hum Brain Mapp* 25:409–423.
- Corbetta M, Shulman GL (2002): Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci* 3:201–215.
- Corbetta M, Kincade JM, Shulman GL (2002): Neural systems for visual orienting and their relationships to spatial working memory. *J Cogn Neurosci* 14:508–523.
- Crummer GC, Walton JP, Wayman JW, Hantz EC, Frisina RD (1994): Neural processing of musical timbre by musicians, non-musicians, and musicians possessing absolute pitch. *J Acoust Soc Am* 95:2720–2727.
- Dapretto M, Bookheimer SY (1999): Form and content: Dissociating syntax and semantics in sentence comprehension. *Neuron* 24:427–432.
- Downar J, Crawley AP, Mikulis DJ, Davis KD (2000): A multimodal cortical network for the detection of changes in the sensory environment. *Nat Neurosci* 3:277–283.
- Fadiga L, Craighero L, Buccino G, Rizzolatti G (2002): Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *Eur J Neurosci* 15:399–402.
- Ferrari PF, Gallese V, Rizzolatti G, Fogassi L (2003): Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *Eur J Neurosci* 17:1703–1714.
- Fiez JA, Balota DA, Raichle ME, Petersen SE (1999): Effects of lexicality, frequency, and spelling-to-sound consistency on the functional anatomy of reading. *Neuron* 24:205–218.
- Fifer RC (1993): Insular stroke causing unilateral auditory processing disorder: Case report. *J Am Acad Audiol* 4:364–369.
- Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC (1995): Improved assessment of significant activation in functional magnetic-resonance-imaging (fMRI)—Use of a cluster-size threshold. *Magn Reson Med* 33:636–647.
- Frankland BW, Cohen AJ (2004): Parsing of melody: Quantification and testing of the local grouping rules of Lerdahl and Jackendoff's generative theory of tonal music. *Music Percept* 21:499–543.
- Friederici AD (2002): Towards a neural basis of auditory sentence processing. *Trends Cogn Sci* 6:78–84.
- Friederici AD, Alter K (2004): Lateralization of auditory language functions: A dynamic dual pathway model. *Brain Lang* 89:267–276.
- Friston KJ, Holmes AP, Poline JB, Grasby PJ, Williams SCR, Frackowiak RS, Turner R (1995): Analysis of fmri time-series revisited. *Neuroimage* 2:45–53.
- Glover GH (1999): Deconvolution of impulse response in event-related BOLD fMRI. *Neuroimage* 9:416–429.
- Griffiths TD (2003): Functional imaging of pitch analysis. *Neurosci Music* 999:40–49.
- Griffiths TD, Warren JD (2002): The planum temporale as a computational hub. *Trends Neurosci* 25:348–353.
- Griffiths TD, Rees A, Witton C, Cross PM, Shalir RA, Green GGR (1997): Spatial and temporal auditory processing deficits following right hemisphere infarction. A psychophysical study. *Brain* 120:785–794.
- Griffiths TD, Buchel C, Frackowiak RSJ, Patterson RD (1998): Analysis of temporal structure in sound by the human brain. *Nat Neurosci* 1:422–427.
- Griffiths TD, Johnsrude I, Dean JL, Green GGR (1999): A common neural substrate for the analysis of pitch and duration pattern in segmented sound? *Neuroreport* 10:3825–3830.
- Habib M, Daquin G, Milandre L, Royere ML, Rey M, Lanteri A, Salamon G, Khalil R (1995): Mutism and auditory agnosia due to bilateral insular damage—role of the insula in human communication. *Neuropsychologia* 33:327–339.
- Halpern AR, Zatorre RJ (1999): When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies. *Cereb Cortex* 9:697–704.
- Halpern AR, Zatorre RJ, Bouffard M, Johnson JA (2004): Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia* 42:1281–1292.
- Haueisen J, Knösche TR (2001): Involuntary motor activity in pianists evoked by music perception. *J Cogn Neurosci* 13:786–792.
- Hickock G, Poeppel D (2000): Towards a functional neuroanatomy of speech perception. *Trends Cogn Sci* 4:131–138.
- Hickock G, Poeppel D (2004): Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition* 92:67–99.
- Hirsh-Pasek K, Nelson DG, Jusczyk PW, Cassidy KW, Druss B, Kennedy L (1987): Clauses are perceptual units for young infants. *Cognition* 26:269–286.

- Holmes AP, Friston KJ (1998): Generalisability, random effects, and population inference. *Neuroimage* 7:754.
- Huettel SA, McCarthy G (2001): The effects of single-trial averaging upon the spatial extent of fMRI activation. *Neuroreport* 12:2411–2416.
- Iidaka T, Matsumoto A, Nogawa J, Yamamoto Y, Sadato N (2006): Frontoparietal network involved in successful retrieval from episodic memory. Spatial and temporal analyses using fMRI and ERP. *Cereb Cortex* 16:1349–1360.
- Imaruoka T, Yanagida T, Miyauchi S (2003): Attentional set for external information activates the right intraparietal area. *Cogn Brain Res* 16:199–209.
- Itoh K, Suwazono S, Arao H, Miyazaki K, Nakada T (2005): Electrophysiological correlates of absolute pitch and relative pitch. *Cereb Cortex* 15:760–769.
- Janata P, Grafton ST (2003): Swinging in the brain: Shared neural substrates for behaviors related to sequencing and music. *Nat Neurosci* 6:682–687.
- Jäncke L, Wüstenberg T, Scheich H, Heinze HJ (2002a): Phonetic perception and the temporal cortex. *Neuroimage* 15:733–746.
- Jäncke L, Wüstenberg T, Schulze K, Heinze HJ (2002b): Asymmetric hemodynamic responses of the human auditory cortex to monaural and binaural stimulation. *Hear Res* 170:166–178.
- Johnson JK, Petsche H, Richter P, vonStein A, Filz O (1996): The dependence of coherence estimates of spontaneous EEG on gender and music training. *Music Percept* 13:563–581.
- Jusczyk PW, Krumhansl CL (1993): Pitch and rhythmic patterns affecting infants sensitivity to musical phrase structure. *J Exp Psychol Hum Percept Perform* 19:627–640.
- Keenan JP, Thangaraj V, Halpern AR, Schlaug G (2001): Absolute pitch and planum temporale. *Neuroimage* 14:1402–1408.
- Knösche TR, Neuhaus C, Haueisen J, Alter K (2003): The role of the planum temporale in the perception of musical phrases. In: *Proceedings of 4th International Conference on Noninvasive Functional Source Imaging (NFSI)*, Chieti, Italy, September 2003.
- Knösche TR, Neuhaus C, Haueisen J, Alter K, Maess B, Witte OW, Friederici AD (2005): Perception of phrase structure in music. *Hum Brain Mapp* 24:259–273.
- Koelsch S, Gunter TC, von Cramon DY, Zysset S, Lohmann G, Friederici AD (2002): Bach speaks: A cortical “language-network” serves the processing of music. *Neuroimage* 17:956–966.
- Koelsch S, Grossmann T, Gunter T C, Hahne A, Schröger E, Friederici AD (2003a): Children processing music: Electric brain responses reveal musical competence and gender differences. *J Cogn Neurosci* 15:683–693.
- Koelsch S, Maess B, Grossmann T, Friederici AD (2003b): Electric brain responses reveal gender differences in music processing. *Neuroreport* 14:709–713.
- Koelsch S, Fritz T, Schulze K, Alsop D, Schlaug G (2005): Adults and children processing music: An fMRI study. *Neuroimage* 25:1068–1076.
- Koelsch S, Fritz T, von Cramon DY, Müller K, Friederici AD (2006): Investigating emotion with music: An fMRI study. *Hum Brain Mapp* 27:239–250.
- Kreutz G, Russ MO, Bongard S, Lanfermann H (2003): Cerebral correlates of music listening. An fMRI-study on the effects of “happy” and “sad” classical music. *Nervenheilkunde* 22:150–156.
- Krumhansl CL, Jusczyk PW (1990): Infants perception of phrase structure in music. *Psychol Sci* 1:70–73.
- Langheim FJP, Callicott JH, Mattay VS, Duyn JH, Weinberger DR (2002): Cortical systems associated with covert music rehearsal. *Neuroimage* 16:901–908.
- Lee DJ, Chen Y, Schlaug G (2003): Corpus callosum: Musician and gender effects. *Neuroreport* 14:205–209.
- Levitin DJ, 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 DJ, 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.
- Liégeois-Chauvel C, Peretz I, Babai M, Laguitton V, Chauvel P (1998): Contribution of different cortical areas in the temporal lobes to music processing. *Brain* 121:1853–1867.
- LoCasto PC, Krebs-Noble D, Gullapalli RP, Burton MW (2004): An fMRI investigation of speech and tone segmentation. *J Cogn Neurosci* 16:1612–1624.
- Lohmann G, Muller K, Bosch V, Mentzel H, Hessler S, Chen L, Zysset S, von Cramon DY (2001): LIPSIA—A new software system for the evaluation of functional magnetic resonance images of the human brain. *Comput Med Imaging Graph* 25: 449–457.
- Lütkenhöner B, Steinsträter O (1998): High-precision neuromagnetic study of the functional organization of the human auditory cortex. *Audiol Neurotol* 3:191–213.
- Maess B, Koelsch S, Gunter TC, Friederici AD (2001): Musical syntax is processed in Broca’s area: An MEG study. *Nat Neurosci* 4:540–545.
- May JK, Assheuer J, Paxinos G (2004): *Atlas of the Human Brain*. San Diego: Elsevier Academic Press.
- Meister IG, Krings T, Foltys H, Boroojerdi B, Muller M, Topper R, Thron A (2004): Playing piano in the mind—An fMRI study on music imagery and performance in pianists. *Cogn Brain Res* 19:219–228.
- Menon V, Levitin DJ, Smith BK, Lembke A, Krasnow BD, Glazer D, Glover GH, McAdams S (2002): Neural correlates of timbre change in harmonic sounds. *Neuroimage* 17:1742–1754.
- Mesulam MM, Mufson EJ (1985): The insula of Reil in man and monkey. Architectonics, connectivity and function. In: Jones EG, Peters A, editors. *Cerebral Cortex*, Vol. 4. New York: Plenum. pp 179–226.
- Meyer M, Steinhauer K, Alter K, Friederici AD, von Cramon DY (2004): Brain activity varies with modulation of dynamic pitch variance in sentence melody. *Brain Lang* 89:277–289.
- Morrison SJ, Demorest SM, Aylward EH, Cramer SC, Maravilla KR (2003): FMRI investigation of cross-cultural music comprehension. *Neuroimage* 20:378–384.
- Müller RA, Kleinhans N, Courchesne E (2001): Broca’s area and the discrimination of frequency transitions: A functional MRI study. *Brain Lang* 76:70–76.
- Nakamura A, Maess B, Knösche TR, Gunter TC, Bach P, Friederici AD (2004): Cooperation of different neuronal systems during hand sign recognition. *Neuroimage* 23:25–34.
- Nan Y, Knösche TR, Friederici AD (2006): The perception of music phrase structure: A cross-cultural ERP study. *Brain Res* 1094: 179–191.
- Neuhaus C, Knösche TR, Friederici AD (2006): Effects of musical expertise and boundary markers on phrase perception in music. *J Cogn Neurosci* 18:472–493.
- Ni W, Constable RT, Mencl WE, Pugh KR, Fulbright RK, Shaywitz SE, Shaywitz BA, Gore JC, Shankweiler D (2000): An event-related neuroimaging study distinguishing form and content in sentence processing. *J Cogn Neurosci* 12:120–133.
- Norris DG (2000): Reduced power multislice MDEFT imaging. *J Magn Reson Imaging* 11:445–451.

- Ohnishi T, Matsuda H, Asada T, Aruga M, Hirakata M, Nishikawa M, Katoh A, Imabayashi E (2001): Functional anatomy of musical perception in musicians. *Cereb Cortex* 11:754–760.
- Oldfield RC (1971): The assessment and analysis of handedness. *Neuropsychologia* 9:97–113.
- Pannekamp A, Toepel U, Alter K, Hahne A, Friederici AD (2005): Prosody-driven sentence processing: An event-related brain potential study. *J Cogn Neurosci* 17:407–421.
- Patel AD, Peretz I, Tramo M, Labreque R (1998): Processing prosodic and musical patterns: A neuropsychological investigation. *Brain Lang* 61:123–144.
- Penfield W, Faulk ME (1955): The insula—Further observations on its function. *Brain* 78:445–470.
- Peretz I, Zatorre RJ (2005): Brain organization for music processing. *Ann Rev Psychol* 56:89–114.
- Platel H, Price C, Baron JC, Wise R, Lambert J, Frackowiak RSJ, Lechevalier B, Eustache F (1997): The structural components of music perception: A functional anatomical study. *Brain* 120:229–243.
- Platel H, Baron JC, Desgranges B, Bernard F, Eustache F (2003): Semantic and episodic memory of music are subserved by distinct neural networks. *Neuroimage* 20:244–256.
- Poeppel D (2001): Pure word deafness and the bilateral processing of the speech code. *Cogn Sci* 25:679–693.
- Poeppel D (2003): The analysis of speech in different temporal integration windows: Cerebral lateralization as ‘asymmetric sampling in time’. *Speech Commun* 41:245–255.
- Poldrack RA, Wagner AD, Prull MW, Desmond JE, Glover GH, Gabrieli JDE (1999): Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage* 10:15–35.
- Ranganath C, Johnson MK, D’Esposito M (2003): Prefrontal activity associated with working memory and episodic long-term memory. *Neuropsychologia* 41:378–389.
- Riemann H (1900): *Vademecum der Phrasierung*. Leipzig: Max-Hesse-Verlag.
- Rizzolatti G, Craighero L (2004): The mirror-neuron system. *Ann Rev Neurosci* 27:169–192.
- Rizzolatti G, Fadiga L, Gallese V, Fogassi L (1996): Premotor cortex and the recognition of motor actions. *Cogn Brain Res* 3:131–141.
- Samson S, Zatorre RJ (1994): Contribution of the right temporal-lobe to musical timbre discrimination. *Neuropsychologia* 32:231–240.
- Satoh M, Takeda K, Nagata K, Hatazawa J, Kuzuhara S (2001): Activated brain regions in musicians during an ensemble: A PET study. *Cogn Brain Res* 12:101–108.
- Satoh M, Takeda K, Nagata K, Hatazawa J, Kuzuhara S (2003): The anterior portion of the bilateral temporal lobes participates in music perception: A positron emission tomography study. *Am J Neuroradiol* 24:1843–1848.
- Schlaug G, Jancke L, Huang YX, Steinmetz H (1995): In-vivo evidence of structural brain asymmetry in musicians. *Science* 267:699–701.
- Schmithorst VJ (2005): Separate cortical networks involved in music perception: Preliminary functional MRI evidence for modularity of music processing. *Neuroimage* 25:444–451.
- Schnyer DM, Verfaellie M, Alexander MP, LaFleche G, Nicholls L, Kaszniak AW (2004): A role for right medial prefrontal cortex in accurate feeling-of-knowing judgments: Evidence from patients with lesions to frontal cortex. *Neuropsychologia* 42:957–966.
- Schnyer DM, Nicholls L, Verfaellie M (2005): The role of VMPC in metamemorial judgments of content retrievability. *J Cogn Neurosci* 17:832–846.
- Schubotz RI, von Cramon DY (2001): Functional organization of the lateral premotor cortex: fMRI reveals different regions activated by the anticipation of object properties, location and speed. *Brain Res* 11:97–112.
- Schubotz RI, Friederici AD, von Cramon DY (2000): Time perception and motor timing: A common cortical and subcortical basis revealed by fMRI. *Neuroimage* 11:1–12.
- Seung Y, Kyong JS, Woo SH, Lee BT, Lee KM (2005): Brain activation during music listening in individuals with or without prior music training. *Neurosci Res* 52:323–329.
- Shah NJ, Marshall NC, Zafiris O, Schwab A, Zilles K, Markowitsch HJ, Fink GR (2001): The neural correlates of person familiarity—A functional magnetic resonance imaging study with clinical implications. *Brain* 124:804–815.
- Shahin A, Bosnyak DJ, Trainor LJ, Roberts LE (2003): Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. *J Neurosci* 23:5545–5552.
- Small DM, Gitelman DR, Gregory MD, Nobre AC, Parrish TB, Mesulam MM (2003): The posterior cingulate and medial prefrontal cortex mediate the anticipatory allocation of spatial attention. *Neuroimage* 18:633–641.
- Spreen O, Benton AL, Fincham RW (1965): Auditory agnosia without multimodaphasia. *Arch Neurol* 13:84–92.
- Steinhauer K, Alter K, Friederici AD (1999): Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nat Neurosci* 2:191–196.
- Sudakov K, McLean PD, Reeves A, Marino R (1971): Unit study of exteroceptive inputs to claustrum in awake sitting squirrel monkeys. *Brain Res* 28:19–34.
- Sundara M, Namasivayam AK, Chen R (2001): Observation-execution matching system for speech: A magnetic stimulation study. *Neuroreport* 12:1341–1344.
- Talairach J, Tournoux P (1988): *Co-planar Stereotaxic Atlas of the Human Brain*. New York: Thieme.
- Thiel CM, Zilles K, Fink GR (2004): Cerebral correlates of alerting, orienting and reorienting of visuospatial attention: An event-related fMRI study. *Neuroimage* 21:318–328.
- Tillmann B, Janata P, Bharucha JJ (2003): Activation of the inferior frontal cortex in musical priming. *Cogn Brain Res* 16:145–161.
- Tramo MJ, Bharucha JJ (1991): Musical priming by the right-hemisphere post-callosotomy. *Neuropsychologia* 29:313–325.
- Ugurbil K, Garwood M, Ellermann J, Hendrich K, Hinke R, Hu XP, et al (1993): Imaging at high magnetic-fields - initial experiences at 4-T. *Magn Reson Q* 9:259–277.
- Vuust P, Roepstorff A, Wallentin M, Moudridsen K, Østergaard L (2006): It don’t mean a thing. Keeping the rhythm during poly-rhythmic tension, activates language areas (BA47). *Neuroimage* 31:832–841.
- Wagner AD, Pare-Blagoev EJ, Clark J, Poldrack RA (2001): Recovering meaning: Left prefrontal cortex guides controlled semantic retrieval. *Neuron* 31:329–338.
- Watkins K, Paus T (2004): Modulation of motor excitability during speech perception: The role of Broca’s area. *J Cogn Neurosci* 16:978–987.
- Watkins KE, Strafella AP, Paus T (2003): Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia* 41:989–994.
- Wheeler ME, Buckner RL (2004): Functional-anatomic correlates of remembering and knowing. *Neuroimage* 21:1337–1349.
- Worsley KJ, Friston KJ (1995): Analysis of fmri time-series revisited—Again. *Neuroimage* 2:173–181.

- Yonelinas AP, Otten LJ, Shaw KN, Rugg MD (2005): Separating the brain regions involved in recollection and familiarity in recognition memory. *J Neurosci* 25:3002–3008.
- Zarahn E, Aguirre G K, DEsposito M (1997): Empirical analyses of BOLD fMRI statistics. I. Spatially unsmoothed data collected under null-hypothesis conditions. *Neuroimage* 5:179–197.
- Zatorre RJ (1997): Cerebral correlates of human auditory processing: Perception of speech and musical sounds. In: Syka J, editor. *Acoustical Signal Processing in the Central Auditory System*. New York: Plenum. pp 453–468.
- Zatorre RJ (2003): Absolute pitch: A model for understanding the influence of genes and development on neural and cognitive function. *Nat Neurosci* 6:692–695.
- Zatorre RJ, Evans AC, Meyer E (1994): Neural mechanisms underlying melodic perception and memory for pitch. *J Neurosci* 14: 1908–1919.
- Zatorre RJ, Halpern AR, Perry DW, Meyer E, Evans AC (1996): Hearing in the mind's ear: A PET investigation of musical imagery and perception. *J Cogn Neurosci* 8:29–46.
- Zatorre RJ, Belin P, Penhune VB (2002): Structure and function of auditory cortex: Music and speech. *Trends Cogn Sci* 6:37–46.