



Investigation of melodic contour processing in the brain using multivariate pattern-based fMRI

Yune-Sang Lee^{a,b,c,*}, Petr Janata^d, Carlton Frost^a, Michael Hanke^{a,b,e}, Richard Granger^{a,b}

^a Dept. of Psychological & Brain Sciences, Dartmouth College, Hanover, NH, USA

^b Center for Cognitive Neuroscience, Dartmouth College, Hanover, NH, USA

^c Neurology Department, University of Pennsylvania, Philadelphia, PA, USA

^d Center for Mind and Brain, U.C. Davis, CA, USA

^e Dept. of Experimental Psychology, Otto-von-Guericke University, Magdeburg, Germany

ARTICLE INFO

Article history:

Received 23 November 2010

Revised 28 January 2011

Accepted 2 February 2011

Available online 21 February 2011

Keywords:

Auditory
Music
Melody
Contour
fMRI
MVPA
Neuroimaging
Multivariate
Spatial
Action perception
Emotion
rSTS
ACC
IPL

ABSTRACT

Music perception generally involves processing the frequency relationships between successive pitches and extraction of the melodic contour. Previous evidence has suggested that the 'ups' and 'downs' of melodic contour are categorically and automatically processed, but knowledge of the brain regions that discriminate different types of contour is limited. Here, we examined melodic contour discrimination using multivariate pattern analysis (MVPA) of fMRI data. Twelve non-musicians were presented with various ascending and descending melodic sequences while being scanned. Whole-brain MVPA was used to identify regions in which the local pattern of activity accurately discriminated between contour categories. We identified three distinct cortical loci: the right superior temporal sulcus (rSTS), the left inferior parietal lobule (IPL), and the anterior cingulate cortex (ACC). These results complement previous findings of melodic processing within the rSTS, and extend our understanding of the way in which abstract auditory sequences are categorized by the human brain.

Published by Elsevier Inc.

Introduction

When listening to music, we effortlessly follow a series of ups and downs between notes in a melody. Moreover, we can easily recognize a well-known musical tune regardless of the key in which it is played. One of the dominant music theories is that musical melodies are encoded via two distinct systems: contour processing, which concerns ups and downs of pitch change irrespective of its exact distance, and interval processing, which analyzes the absolute or relative distance from one note to another (Dowling, 1978; Dowling and Fujitani, 1971; Peretz, 1990; Peretz and Babai, 1992). Although interval processing is important for establishing the tonality (key) of a musical passage, considerable evidence has suggested that contour processing provides an essential basis for melody recognition (Dowling, 1978; Dowling and Fujitani, 1971; Edworthy,

1985; Hebert and Peretz, 1997; Peretz and Babai, 1992). For example, Dowling and Fujitani (1971) investigated the role of both interval and contour in melody recognition. In one condition, subjects were required to detect subtle changes in interval size between reference and target melodies that were played in either the same or different keys. The results showed that performance was worse when the target melodies were presented in different keys than when presented in the original key. In another experiment, subjects were required to detect changes in the contours of target melodies that were played in the same or different keys. For this task, subjects' performance was robust across transposition to different keys. The experiments suggest that contour is a defining feature of melodies, and thereby a potent cue to the identity of musical pieces, for instance the short note sequences that form recognizable 'hooks' and 'themes'. Further behavioral studies revealed that both infants and musically naïve adults were able to detect contour but not interval changes in unfamiliar melodies across changes of key (Trainor and Trehub, 1992; Trainor and Trehub, 1994). Together, these studies have suggested that contour may be a more fundamental attribute for melody recognition than interval size.

* Corresponding author at: Center for functional neuroimaging, Neurology department, 3W. Gates Bldg., 3400 Spruce St., Philadelphia, PA, 19104-4283, USA. Fax: +1 215 349 8260.

E-mail address: yslee@mail.med.upenn.edu (Y.-S. Lee).

A melodic contour can be categorically parsed into its minimal units, namely ups and downs. An influential lesion study (Johnsrude et al., 2000) revealed that patients with a lesion in the right superior temporal lobe could reliably distinguish whether two pitches were the same or different, but were impaired when judging whether the second note was higher or lower than the first. This partial impairment clearly indicates that there exist neural substrates that provide information about the directionality of successive pitches beyond the tonotopic organization of the primary auditory cortex (A1).

Patterson et al. (2002) showed this to be the case using spectrally-matched auditory stimuli. In this sophisticated fMRI study, contrast of melodies (either random or diatonic scale) versus fixed-pitch sound revealed activity in the right superior temporal region whereas contrast of fixed-pitch sound versus non-pitch sound revealed activity mostly within the Heschl's gyrus. This finding suggests a hierarchical organization of complex pitch processing. Relatedly, Stewart et al. (2008) found that detection of 'local' interval violations in 4-pitch sequences primarily recruited regions of the right posterior superior temporal sulcus (STS), whereas detection of more 'global' contour violations preferentially recruited the left posterior STS.

The present fMRI study sought to better understand 'where' and 'how' melodic contour information is represented along the auditory pathway. To this end, we chose a multivariate pattern-based fMRI analysis (MVPA) approach. Although the subtraction logic of standard neuroimaging techniques has offered a way to highlight the network of areas associated with melody processing (Hyde et al., 2008, 2011; Janata et al., 2002a, 2002b; Patterson et al., 2002; Platel et al., 1997; Stewart et al., 2008; Warren et al., 2003; Warren and Griffiths, 2003; Zatorre et al., 1994, 1996), the method suffers if the objective is to differentiate stimulus categories and all of the voxels in a region are modulated by one category or the other. This limitation led us to believe that the conventional neuroimaging approach may not be suited to address our question. Instead, we expected that 'ups' and 'downs' of contour category may be found by measuring differential voxel patterns across the stimuli using MVPA.

One MVPA method, the searchlight analysis developed by Kriegeskorte et al. (2006), has been proven to effectively delineate brain regions that are inherently invisible to the standard fMRI analysis method (Kriegeskorte et al., 2006; Raizada et al., 2010; for a tutorial review, see Pereira et al., 2009). With this approach, we examined each location of the brain to identify areas that may categorize between ascending and descending melodies in a non-musician group.

Additionally, the identical analysis method was applied to explore the brain regions that may be involved in mode processing by collapsing the same set of stimuli into major and minor categories.

Following the fMRI experiment, two behavioral experiments were conducted 1) to construct a perceptual similarity space with regards to all the melodies and 2) to evaluate the emotional valence (e.g., happy vs. sad) aroused by each melodic sequence. These behavioral experiments would further support the neural findings and provide interpretability of the neural findings in relation to the observed behavior.

Materials and methods

Subjects

Subjects were 12 healthy right-handed volunteers (7 male; average age = 20.4; average musical training = 5.7 years), none of whom majored in music nor had participated in professional or semi-professional music activities (e.g., playing in an orchestra or a rock band). No subjects had absolute pitch. Consent forms were obtained from all subjects as approved by the Committee for the Protection of Human Subjects at Dartmouth College.

Stimuli

Twenty short melodic sequences consisting of five piano tones in the middle octave range were generated using the MIDI sequence tool in Apple's Garage Band software and exported to .wav format. All stimuli were matched in duration (2.5 s, 500 ms per each note), sampling rate (44.1 kHz, 16-bit, Stereo), and volume intensity using SoundForge 9.0 (Sony, Japan) and Matlab 2009b (Mathworks Inc, Natick, MA, USA). A 2 × 2 design was employed with Mode (major, minor) in one dimension and Contour (ascending, descending) in another dimension, creating four categories of stimuli, each of which contained five different melody exemplars whose slopes were systematically varied (Fig. 1). In addition to these 20 stimuli, melodies with a third type of contour comprising both upward and downward pitch changes were created to be used as catch trials during the scans (Supplementary Fig. 1).

fMRI scanning

A slow event-related design was employed with an 8 s inter-stimulus interval (ISI) in eight runs (44 trials per run). Fixation crosses were displayed during runs. Scanning was conducted on a 3 T Philips Intera Achieva whole body scanner (Philips Medical Systems, Best, the Netherlands) at the Dartmouth College Brain Imaging Center. The parameters of standard echo-planar imaging (EPI) sequences are as follows: TR = 2000 ms, TE = 35 ms, FOV = 240 × 240 mm, 30 slices, voxel size = 3 × 3 × 3 mm, and inter-slice interval = 0.5 mm, sequential axial acquisition. A high-resolution T1-weighted MPRAGE scan



Fig. 1. Staff view of the 20 melody stimuli. a. Diatonic scale, b. 7th scale, c. Arpeggio scale, d. 5th scale, e. Wide arpeggio scale. All melodies were anchored to the note of C (261.63 Hz) in the middle octave range (261.63 Hz). The tempo of each melodic sequence was 120 bpm.

(voxel size = $1 \times 1 \times 1$ mm) was acquired at the end of the scan. Stimuli were delivered binaurally using high-fidelity MR-compatible headphones (OPTIME 1, MR confon, Germany).

Experimental procedures

fMRI experiment

During the scan, subjects heard a series of melodies while fixating the cross on the screen. Each melody was presented twice per run (a total of 40 melodies per run) every 8 s and the order was randomized across the runs. Occasionally, a catch trial melody (Supplementary Fig. 1) was presented to monitor subjects' alertness (a total of 4 melodies) for which subjects were instructed to press a button to indicate when they perceived a change in contour within a particular melody.

Happiness rating

Following fMRI scans, happiness rating was measured on a separate day. In a quiet behavioral testing room, stimuli were presented via noise-canceling headphones (Quiet Comfort acoustic noise-canceling headphones, Bose, USA) and subjects were instructed to report how happy each melody sounded using a Likert-type scale from 1 (*very sad*) to 7 (*very happy*).

Similarity distance measurement

For another post-fMRI experiment, similarity distance among the melodies was measured. In a quiet behavioral testing room, consecutive pairs of sequences consisting of the stimuli from the fMRI experiment were presented via noise-canceling headphones (Quiet Comfort acoustic noise-canceling headphones, Bose, USA) and subjects were asked to indicate how similar each pair of melodies (400 pairs, 20×20) sounded using a Likert-type scale from 1 (*not at all similar*) to 7 (*exactly alike*). Subjects were encouraged to use the full scale. The full list comprised the set of all possible pairings, presented over the course of two half-hour sessions.

MVPA methods

fMRI data were preprocessed using the SPM5 software package (Institute of Neurology, London, UK) and MATLAB 2009b (Mathworks Inc., Natick, MA, USA). All images were realigned to the first EPI to correct movement artifacts, and then spatially normalized into Montreal Neurological Institute (MNI) standard stereotactic space (e.g., ICBM152 EPI template) with their preserved original voxel size ($3 \text{ mm} \times 3 \text{ mm} \times 3 \text{ mm}$).

Classification on the contour (ascending vs. descending)

After preprocessing, fMRI time-courses of all voxels were extracted from unsmoothed images. Subsequently, these raw signals were high-pass filtered with a 300 s cut-off to remove slow drifts caused by the scanner, and standardized across entire runs using z-score to normalize intensity differences among runs. To guard against a confounding signal from different stimulus onsets, a signal solely generated by each stimulus (i.e., corresponding to time points 4, 6, and 8 s after the onset of the stimulus) was acquired from voxels. The neural signals corresponding to ascending and descending categories were vectorized to be submitted to a classifier. For the binary classifier, we used the Lagrangian Support Vector Machine algorithm (Mangasaian and Musicant, 2001). The classifier was initially trained by a subset of datasets (training set) and applied to the remaining datasets (testing set). For the purpose of validating results, signals from six scanning runs served as a training set and two runs served as a testing set, resulting in 4-fold cross-validation. The classification was performed approximately 50,000 times at every searchlight sphere (a radius size consisting of two neighboring voxels, maximum 33 voxels). The percent correct result for each classification test was averaged across the four training/testing combinations and stored in

each voxel (the center voxel of a sphere) of an output image for each subject. These output images of all subjects were submitted to a second-level random effect (Raizada et al., 2010; Walther et al., 2009; Stokes et al., 2009) using SPM such that the average accuracy of classification test for each voxel was compared to chance (50%) and the group *t*-map containing the corresponding *t*-value for each voxel was generated.

Classification on the mode (major vs. minor)

Procedures were identical except that the corresponding neural signals were chosen based upon mode (major and minor).

Results

Behavioral results (similarity distance matrix)

Similarity data were acquired from 7 out of 12 subjects who had previously participated in the fMRI experiment, compiled in a square symmetrical matrix format, and analyzed using SPSS v. 17.0 (Chicago, IL), generating 2-dimensional Euclidean-distance plots both within and across subjects with 5-stress convergence of .001. The multi-dimensional scaling (MDS) structure revealed that the primary dimension of clustering among the sequences was contour (Fig. 2a). More specifically, it was observed that melodies within the same contour tended to cluster together more than melodies across different contours. The result confirmed our notion that the contour is categorically perceived. Another dimension captured the degree of slope across the stimuli, in that high-slope melodies such as wide arpeggio and 5th melodies tended to cluster together and low-slope melodies such as diatonic and 7th melodies tended to cluster together (Fig. 2a). Together the first two dimensions explained 82% of the variance in the similarity judgments. When added, a 3rd dimension, which appears to be weakly associated with mode, increased the percentage of explained variance by only 2% (see Supplementary Figs. 2a and b).

Behavioral results (happiness rating)

A one-way repeated-measures ANOVA was performed on the average happiness ratings across the four melodic categories. The results revealed that there was a significant difference among the four categories, $F_{(3,33)} = 19.99$, $P < 0.05$ (Fig. 2b). In line with the previous report (Collier and Hubbard, 2001), there was a main effect of contour such that ascending melodies sounded happier than descending melodies irrespective of mode, $t_{11} = 5.58$, $P < 0.05$. Likewise, there was a main effect of mode such that major melodies sounded happier than minor melodies irrespective of contour, $t_{11} = 4.90$, $P < 0.05$. There was no interaction between contour and mode (*n.s.*).

fMRI results

Ascending vs. descending (contour)

The searchlight analysis revealed three distinct brain regions that reliably categorized between ascending and descending melodies (p (*cluster-size corrected*) < 0.05 in combination with p (*uncorrected*) < 0.005) (Fig. 3 and Table 1). Among the areas, a part of the right STS (x, y, z : 51, -18, -7) exhibited the most differential neural pattern between the categories, $t_{11} = 7.71$, confirming the previous findings that melodic processing is mainly mediated by the right superior temporal region (Hyde et al., 2008; Johnsrude et al., 2000; Warrier and Zatorre, 2004; Zatorre, 1985; Zatorre et al., 1994). The IPL on the contralateral left hemisphere also displayed a categorical neural pattern in response to ascending and descending melodies. Within this region, the most robust local pattern ($t_{11} = 5.59$) was observed in the intraparietal sulcus (x, y, z : -48, -36, 39). Finally, the ACC (x, y, z : 3, 21, 28) in the frontal lobe was found to discriminate between

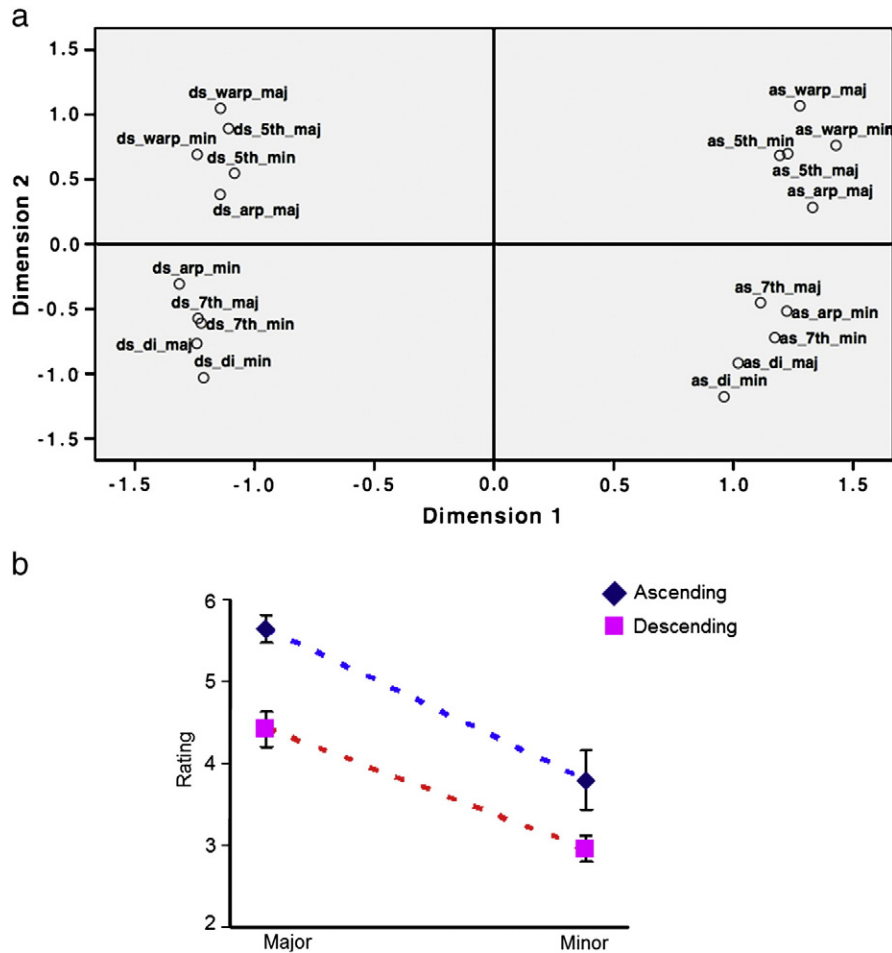


Fig. 2. a. The similarity structure in 2D projections among all pairwise (20 x 20) melody comparisons. The horizontal axis captures the distinction in contour and the vertical axis captures the variance of the slopes among melodies. Together, they account for a total of 82% of variances. Abbreviations: as = ascending, ds = descending, di = diatonic, arp = arpeggio, warp = wide arpeggio, maj = major, min = minor. b. Happiness ratings for the four melody categories. The x-axis depicts 2 different melodic categories by mode, lines by contour, and the y-axis depicts the rating between 1 and 7. There was significant difference among the four melody types in their emotional content ($F_{(3,33)} = 19.99, P < 0.05$).

ascending and descending melodies. Subsequently, ROIs (Regions of Interest) were extracted for verification of overall accuracy within each of the identified areas. The number of voxels within the ROIs was 133 (rSTS), 183 (IIPL), and 233 (ACC) and the fMRI intensity in each ROI was submitted to another set of classification tests. The overall accuracies of classification tests were 52.1% ($s.e = 0.6, t_{11} = 3.5, P = 0.005$), 51.7% ($s.e = 0.6, t_{11} = 3.1, P = 0.01$), 52.3% ($s.e = 0.4, t_{11} = 5.3, P = 0.0003$) in the rSTS, IIPL, and ACC respectively.

While significant, the observed percent accuracies were somewhat low. In order to validate the overall accuracy within each of those areas, simulation using Monte Carlo shuffling was performed. To this end, fMRI signals within each area were randomly assigned to ascending or descending categories in a training set and submitted to the classifier. This classifier was then applied to predict the categories that correctly corresponded to fMRI signals in the remaining testing set. This was done across the 4-fold cross-validation sets. The Monte Carlo shuffling with 1000 iterations was plotted and compared to the observed accuracy (Fig. 4). It was revealed that the observed accuracy was indeed significantly above the chance level ($P < 0.001$) for all three ROIs confirming that each area can distinguish between ascending and descending categories.

Major vs. minor (mode)

Although our primary focus was an investigation of contour processing, we performed a second searchlight analysis to identify the brain regions involved in implicit categorization of melodies by

major/minor mode during the contour task. The analysis did not yield any significant voxels for this classification even at an extremely relaxed significance threshold ($p(\text{uncorrected}) < 0.01$).

Discussion

This study sought to identify neural structures underlying melodic contour processing. More specifically, we searched for sets of voxels that can distinguish between ascending and descending melodic sequences. Using a whole brain searchlight method, we found that three distinct areas, namely, the right STS, left IPL, and ACC produced differential neural patterns in response to ascending and descending contour categories (Fig. 3). By contrast, a subsequent searchlight analysis with respect to mode (major vs. minor) did not find any significant voxels. Given that our behavioral results showed that both mode and contour were separable by emotional valence, these null results may suggest that the categorical neural responses that were observed during the contour task may have more to do with pitch sequence processing rather than higher-order emotional processing. Nonetheless, the possibility of categorization by emotion cannot be fully discounted by the results as the contribution of mode to an emotional differentiation of the melodies could have been disregarded during the contour-detection task. Further, both the rSTS and ACC have been implicated in emotion literature, leaving open the possibility of functional heterogeneity within those regions.

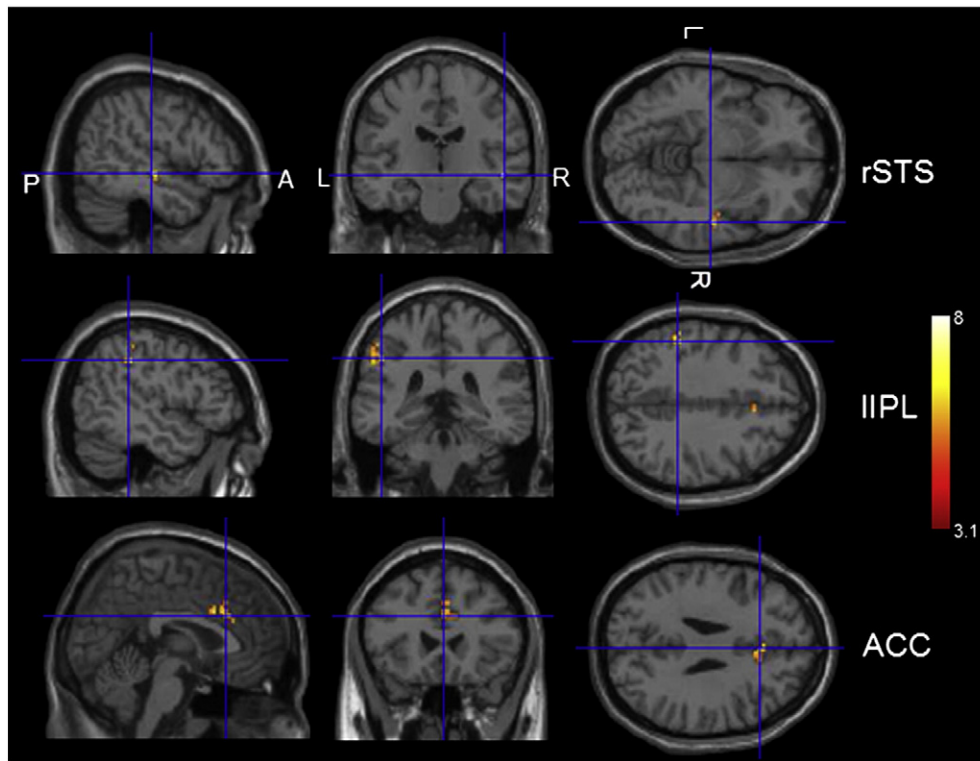


Fig. 3. Brain regions that distinguish between ascending and descending melodic sequences ($P(\text{cluster-size corrected}) < 0.05$ in combination with $P(\text{uncorrected}) < 0.005$). The color scale indicates separability t -value from the second level group analysis of individual searchlight analysis. Top: right superior temporal sulcus, middle: left inferior parietal lobule, bottom: anterior cingulate cortex.

In line with the neural findings, the similarity distance matrix revealed that ascending and descending melodies were indeed categorically divided in perceptual space (Fig. 2a). As was expected, GLM analysis yielded no significant voxels except for the melodies vs. resting period comparison, which mainly showed bilateral activation of the auditory cortices (see supplemental material).

In the field of cognitive neuroscience of music, the neural signature of contour processing has been primarily investigated in EEG (electroencephalography) studies (Fujioka et al., 2004; Paavilainen et al., 1998; Saarinen et al., 1992; Schiavetto et al., 1999; Tervaniemi et al., 1994; Trainor et al., 2002). For example, Trainor et al. (2002) found that a contour shift from ascending to descending melodies elicited a MMN (mismatch negativity) in musically untrained subjects, suggesting that melodic contour information might be categorized automatically by the brain even in the absence of attention. Moreover, a subsequent MEG (magnetoencephalography) study (Fujioka et al., 2004) extended the previous findings by demonstrating that the MMN was more pronounced in musicians than in non-musicians when a contour change occurred, indicating that automatic contour processing can be sharpened by musical experience.

While EEG (MEG) studies have been conducted in the temporal domain, neuroimaging work has made a substantial contribution to create the spatial map of melody processing in the brain (Hyde et al.,

2008, 2011; Janata et al., 2002a, 2002b; Patterson et al., 2002; Platel et al., 1997; Stewart et al., 2008; Warren et al., 2003; Warren and Griffiths, 2003; Zatorre et al., 1994, 1996). A number of neuroimaging studies have revealed that melodies tend to evoke activation in regions of the right hemisphere, including superior temporal and inferior frontal lobes (Hyde et al., 2008, 2011; Janata et al., 2002a; Patterson et al., 2002; Zatorre et al., 1994). For example, an early PET (Positron Emission Tomography) study by Zatorre et al. (1994) showed that the right superior temporal sulcus was more activated when listening to a melody than to a noise burst matched in amplitude envelope. Hyde et al. (2008) showed that the right planum temporale was parametrically modulated by the degree of pitch distance in melodic sequences whereas the left planum temporale was not responsive until the pitch distance was increased up to 200 cents between adjacent notes. In a more recent study by Hyde et al. (2011), right inferior frontal gyrus showed deactivation and reduced functional connectivity with the auditory cortex in amusics who were impaired in pitch processing when compared to normal subjects.

Nevertheless, as was discussed in the Introduction, the conventional neuroimaging paradigm may be limited when directly comparing different melody stimuli. Notably, our study showed that the conventional approach was blind to the brain regions that generate differential patterns in response to different contours of melodies that were matched in other physical characteristics such as tempo and duration (see supplemental material). While significant, the classification test indicated that the distinction in each area was quite subtle. This may have been due to normalizing subjects' brains into the standard MNI template, which inevitably disregards anatomical differences across individuals. Our aim, however, was to identify melodic contour modules at the group level, not to test classification results in each subject's pre-defined ROIs. We attempted to validate the group results by simulating classification accuracies that were derived from randomly shuffled data with 1000 iterations. The Monte Carlo results for each area confirmed that the observed

Table 1
Cortical loci involved in melodic contour processing.

Region name	HEM	BA	MNI coordinates			T-value	Cluster size
			x	y	z		
Superior temporal sulcus	R	22	51	-18	-7	7.71	133
Inferior parietal lobule	L	40	-48	-36	39	5.59	183
Anterior cingulate	R	32	3	24	28	4.66	233

Table lists significant areas ($p(\text{cluster-size corrected}) < 0.05$).
HEM: hemisphere; BA: approximate Brodmann area.

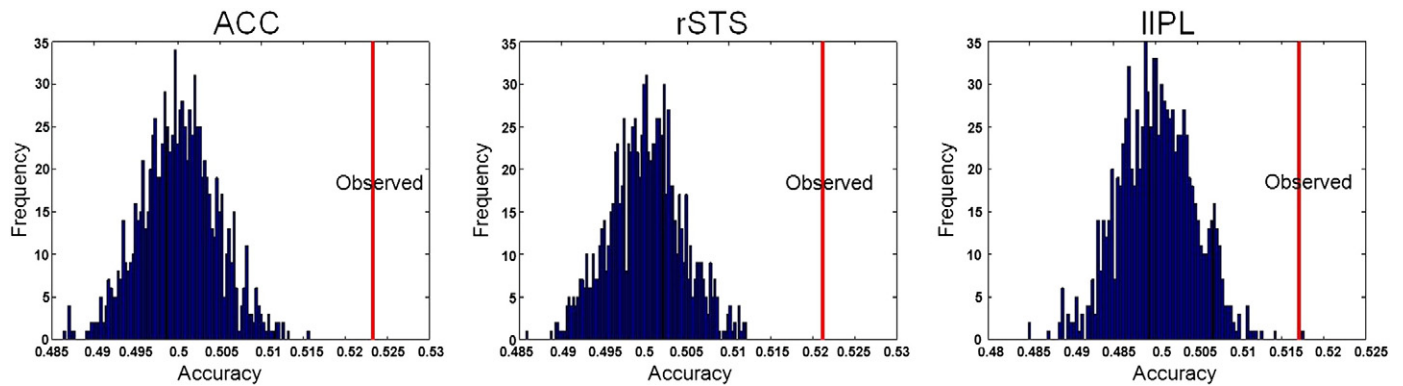


Fig. 4. Results of Monte Carlo shuffling with 1000 iterations in each ROI found by whole brain searchlight analysis. The vertical red line indicates accuracy that was actually observed in the data. None of the Monte Carlo simulation results matched or exceeded the observed accuracy ($P < 0.001$), confirming that the classifications did not merely occur by chance and were significant in all areas.

accuracies from actual data were robust as nearly none of the iterations exceeded the accuracy by random simulation (Fig. 4). Thus, we were convinced that the three identified areas were able to significantly distinguish ascending and descending melodic contours.

The role of the rSTS

Regions along the STS have been reported to play a role in the processing of pitch change and melody (Griffiths et al., 1998; Stewart et al., 2008; Warren et al., 2003; Zatorre et al., 1994), as have slightly more superior areas along the superior temporal gyrus (Patterson et al., 2002; Warren et al., 2003; Warren and Griffiths, 2003). Although a principle appears to hold, whereby more abstract and complex pitch and sequential representations (melodies) are processed farther away from the primary auditory cortex (Griffiths et al., 1998; Patterson et al., 2002; Warren and Griffiths, 2003), considerable variation has been observed in terms of the rostrocaudal location of reported activation maxima, and even hemispheric lateralization (Stewart et al., 2008). The STS focus we observed fell within the middle of the reported rostrocaudal range, ventral to Heschl's gyrus.

The right hemisphere lateralization of the STS focus we observed was also consistent with the result from a lesion study in which individuals with anterior temporal lobe resections that encroached on Heschl's gyrus were impaired in judging whether a note was higher or lower than a preceding note (Johnsrude et al., 2000). Pitch direction judgments may be regarded as a building block for the perception of melodic contours, though behavioral studies suggest that different mechanisms govern pitch judgments about pairs of notes in isolation (intervals) and those presented in longer sequences (Cuddy and Cohen, 1976; Lee et al., submitted). Despite studies demonstrating hierarchical and spatially separate mechanisms for pitch processing in the superior portions of the lateral temporal lobes (Patterson et al., 2002; Warren et al., 2003), the exact distributions of representational mechanisms and melody representations within the STG/STS region along their entire rostrocaudal extent remain to be determined.

It has been argued that the superior temporal region in each hemisphere is specialized for processing different aspects of sounds such that the right superior temporal region is involved in spatial processing (e.g., pitch distance) whereas the left superior temporal region is involved in temporal processing (Hyde et al., 2008; Griffiths and Warren, 2004; Peretz and Zatorre, 2005; Warrier and Zatorre, 2004). Our finding is consistent with the prevailing notion of the functional asymmetry of left and right temporal lobes insofar as only the right superior temporal region emerged to distinguish the spatially orthogonal melodic sequences that were matched in their temporal characteristics. Future MVPA studies may examine whether the left STS can distinguish melodies that differ in the temporal domain.

The right STS has been also implicated in studies examining emotional prosody in music and languages (Ethofer et al., 2009; Grandjean et al., 2005; for review, see Schirmer and Kotz, 2006). Notably, a recent MVPA study demonstrated that they were able to decode emotional information in the voice prosody within the superior temporal cortices (Ethofer et al., 2009). Thus, it is plausible that the rSTS may play a more general role in differentiating pitch contours. Whether any such generalized contour differentiation or categorization process is driven primarily by contours that have affective implications, such as the ascending and descending contours in our experiment, remains to be conclusively established.

The role of the left IPL

It is well known that the parietal lobe is one of the key regions for the dorsal visual 'where' pathway (Mishkin et al., 1983). While numerous studies have shown that this area is activated by visual spatial tasks, the posterior parietal lobe is also known to be involved in multi-modal processing (Schroeder and Foxe, 2002). A number of neurophysiological studies have revealed that the parietal lobe receives sensory inputs from visual, auditory, and tactile sensory cortices (see review by Cohen, 2009). The parietal lobe appears to be involved in symbolic mapping, such as for phonemes (Shaywitz et al., 2003; Temple, 2002), numbers (Kadosh et al., 2010), and musical notes (Schon et al., 2002). The IPL, however, was also implicated in the previous music fMRI studies. For instance, Janata et al. (2002b) reported that a part of the left IPL ($x, y, z: -56, -34, 40$) was more activated when subjects actively attended to melodies in polyphonic music than when they passively listened to acoustically matched stimuli. The locus in the previous study was 8.3 mm distant from the IPL region ($x, y, z: -48, -36, 39$) that we identified here. In a recent music fMRI study, Foster and Zatorre (2010) showed that bilateral IPS was more activated during a relative pitch judgment task compared to a passive melody listening task. They suggested that the spatial transformation on the relative size of pitch distance may be mediated by this region. However, the area ($x, y, z: -40, -60, 54$) was quite distant from the IPS area identified here. Klostermann et al. (2009) reported that the right IPL was recruited during successful retrieval of melodies. Together, these studies suggest that the IPL plays an important role in the processing of melodic sequences. Further research is warranted to examine the specific role of the IPL in each hemisphere in melodic processing.

The role of the ACC

The ACC is implicated in many aspects of cognition, including the processing of musical structure and affective responses to music. It is known to play a key role in attention and behavioral

monitoring (Crottaz-Herbette and Menon, 2006). However, it is quite unlikely that subjects paid differential attention to ascending and descending melodies during the task, so it is unlikely that this domain-general role could account for the discrimination of contour direction.

Among other possibilities, one tenable explanation is in relation to the 'action perception' theory. Converging evidence suggests that the ACC plays a key role in action perception (Botvinick et al., 1999; Calvo-Merino et al., 2005; Posner et al., 1988; also for overview, see Posner, 2004). On the musical side, the long-standing idea that melodic contours have affective implications (Kivy, 1980; Langer, 1957) has recently been echoed in relation to the perception/action coupling in music and the brain's mirror–neuron system (Molnar-Szakacs and Overy, 2006). The two categories of melodies we used (ascending and descending) had affective connotations and can be taken to have strong action implications. Both of these aspects could be expected to drive ACC activity. Indeed, the univariate analysis between all the melody conditions and resting periods showed that the supplementary motor area along with a part of the ACC became more active (Supplementary Fig. 3).

Differences with previous studies of melody processing

Activation of the right lateral inferior frontal lobe has been reported in the neuroimaging literature of melody processing (Hyde et al., 2011; Zatorre et al., 1994). For example, Zatorre and colleagues (1994) demonstrated that several parts of the right inferior frontal lobe were more activated during the condition when subjects were required to judge whether the second (or the last) note was higher or lower than the first note in the presented melodies than the condition when subjects were instructed to passively listen to the melody stimuli. Our analyses did not reveal discrimination of melodic contours in these areas. However, several important differences between the previous and current study may account for the failure of inferior frontal activity here. First, this could be due to the degree to which working memory load is imposed on the tasks. In the previous study (Zatorre et al., 1994), subjects were required to respond after the completion of the 16-note complex melody, while in the present study subjects were required to respond as soon as they detected the violation of contour of much simpler 5-note melodies. Thus, it is likely that simple change monitoring rather than working-memory mechanisms came into play to succeed in the task. Secondly, in the previous study, it is plausible that subjects may have employed different strategies, such as discrimination of absolute pitch, for the given tasks. By contrast, in the present study, pitch discrimination could not come into play to successfully achieve the goal of the task since every note in the stimuli was different in its absolute pitch; attending to the on-going update of contour information, irrespective of absolute pitch differences, was more necessary in the task of our study. Further work is needed to understand more specific contributions of the ACC as well as other prefrontal structures to melodic processing.

Conclusion

The present fMRI study examined melodic contour processing in the brain using multivariate pattern-based analysis (MVPA). The results revealed that three distinct brain regions, the rSTS, IIPPL, and ACC, displayed differential neural activity in response to ascending and descending melodic contours. Several complementary behavioral (similarity matrix, emotional rating) as well as functional (major vs. minor mode classification, Monte Carlo simulation) data analyses further supported the conclusion that these areas play an important role in melodic contour processing. Our findings are in accordance with the prevailing model in which abstract auditory cues are hierarchically processed along the downstream auditory pathway. Moreover, these results support previous findings of melodic proces-

ing within the rSTS, and extend our understanding of the neural basis underlying melodic contour categorization. Future studies should examine the cross-talk of the three identified areas in the current study and how they interact in music processing using appropriate experimental designs and better-suited methods (e.g., functional connectivity on the functional data or diffusion tensor imaging study on the anatomical data).

Supplementary materials related to this article can be found online at doi:10.1016/j.neuroimage.2011.02.006.

Acknowledgment

The authors would like to thank Jim Haxby and Yaroslav Halchenko for their helpful advice on data analysis. Special thanks go to Becca Loucks and Carlton P. Frost V for their help on manuscript proofreading.

References

- Botvinick, M., Nystrom, L.E., Fissell, K., Carter, C.S., Cohen, J.D., 1999. Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature* 402 (6758), 179–181.
- Calvo-Merino, B., Glaser, D.E., Grezes, J., Passingham, R.E., Haggard, P., 2005. Action observation and acquired motor skills: an fMRI study with expert dancers. *Cereb. Cortex* 15 (8), 1243–1249.
- Cohen, Y.E., 2009. Multimodal activity in the parietal cortex. *Hear. Res.* 258 (1–2), 100–105.
- Collier, W.G., Hubbard, T.L., 2001. Musical scales and evaluations of happiness and awkwardness: effects of pitch, direction, and scale mode. *Am. J. Psychol.* 114 (3), 355–375.
- Crottaz-Herbette, S., Menon, V., 2006. Where and when the anterior cingulate cortex modulates attentional response: combined fMRI and ERP evidence. *J. Cogn. Neurosci.* 18 (5), 766–780.
- Cuddy, L.L., Cohen, A.J., 1976. Recognition of transposed melodic sequences. *Q. J. Exp. Psychol.* 28, 255–270 May.
- Dowling, W.J., 1978. Scale and contour – 2 components of a theory of memory for melodies. *Psychol. Rev.* 85 (4), 341–354.
- Dowling, W.J., Fujitani, D.S., 1971. Contour, interval, and pitch recognition in memory for melodies. *J. Acoust. Soc. Am.* 49 (2, Pt. 2), 524–531.
- Edworthy, J., 1985. Interval and contour in melody processing. *Music Percept.* 2 (3), 375–388.
- Ethofer, T., Van De Ville, D., Scherer, K., Vuilleumier, P., 2009. Decoding of emotional information in voice-sensitive cortices. *Curr. Biol.* 19 (12), 1028–1033.
- Foster, N.E., Zatorre, R.J., 2010. A role for the intraparietal sulcus in transforming musical pitch information. *Cereb. Cortex* 20 (6), 1350–1359.
- Fujioka, T., Trainor, L.J., Ross, B., Kakigi, R., Pantev, C., 2004. Musical training enhances automatic encoding of melodic contour and interval structure. *J. Cogn. Neurosci.* 16 (6), 1010–1021.
- Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M.L., Scherer, K.R., Vuilleumier, P., 2005. The voices of wrath: brain responses to angry prosody in meaningless speech. *Nat. Neurosci.* 8 (2), 145–146.
- Griffiths, T.D., Warren, J.D., 2004. What is an auditory object? *Nat. Rev. Neurosci.* 5 (11), 887–892.
- Griffiths, T.D., Buchel, C., Frackowiak, R.S., Patterson, R.D., 1998. Analysis of temporal structure in sound by the human brain. *Nat. Neurosci.* 1 (5), 422–427.
- Hebert, S., Peretz, I., 1997. Recognition of music in long-term memory: are melodic and temporal patterns equal partners? *Mem. Cogn.* 25 (4), 518–533.
- Hyde, K.L., Peretz, I., Zatorre, R.J., 2008. Evidence for the role of the right auditory cortex in fine pitch resolution. *Neuropsychologia* 46 (2), 632–639.
- Hyde, K.L., Zatorre, R.J., Peretz, I., 2011. Functional MRI evidence of an abnormal neural network for pitch processing in congenital amusia. *Cereb. Cortex* 21 (2), 292–299.
- Janata, P., Birk, J.L., Van Horn, J.D., Leman, M., Tillmann, B., Bharucha, J.J., 2002a. The cortical topography of tonal structures underlying Western music. *Science* 298 (5601), 2167–2170.
- Janata, P., Tillmann, B., Bharucha, J.J., 2002b. Listening to polyphonic music recruits domain-general attention and working memory circuits. *Cogn. Affect. Behav. Neurosci.* 2 (2), 121–140.
- Johnsrude, I.S., Penhune, V.B., Zatorre, R.J., 2000. Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain* 123 (Pt 1), 155–163.
- Kadosh, R.C., Muggleton, N., Silvanto, J., Walsh, V., 2010. Double dissociation of format-dependent and number-specific neurons in human parietal cortex. *Cereb. Cortex* 20 (9), 2166–2171.
- Kivy, P., 1980. *The Corded Shell: Reflections on Musical Expression*. Princeton University Press, Princeton.
- Klostermann, E.C., Loui, P., Shimamura, A.P., 2009. Activation of right parietal cortex during memory retrieval of nonlinguistic auditory stimuli. *Cogn. Affect. Behav. Neurosci.* 9 (3), 242–248.
- Kriegeskorte, N., Goebel, R., Bandettini, P., 2006. Information-based functional brain mapping. *Proc. Natl. Acad. Sci. USA* 103 (10), 3863–3868.
- Langer, S.K., 1957. *Philosophy in a New Key: A Study in the Symbolism of Reason, Rite, and Art*, 6th ed. New American Library, Cambridge.

- Mangasaian, O.L., Musicant, D.R., 2001. Lagrangian support vector machines. *J. Mach. Learn. Res.* 1 (3), 161–177.
- Mishkin, M., Ungerleider, L.G., Macko, K.A., 1983. Object vision and spatial vision – 2 cortical pathways. *Trends Neurosci.* 6 (10), 414–417.
- Molnar-Szakacs, I., Overy, K., 2006. Music and mirror neurons: from motion to 'e' motion. *Soc. Cogn. Affect. Neurosci.* 1 (3), 235–241.
- Paavilainen, P., Jaramillo, M., Naatanen, R., 1998. Binaural information can converge in abstract memory traces. *Psychophysiology* 35 (5), 483–487.
- Patterson, R.D., Uppenkamp, S., Johnsrude, I.S., Griffiths, T.D., 2002. The processing of temporal pitch and melody information in auditory cortex. *Neuron* 36 (4), 767–776.
- Pereira, F., Mitchell, T., Botvinick, M., 2009. Machine learning classifiers and fMRI: a tutorial overview. *Neuroimage* 45 (1), S199–S209.
- Peretz, I., 1990. Processing of local and global musical information by unilateral brain-damaged patients. *Brain* 113, 1185–1205.
- Peretz, I., Babai, M., 1992. The role of contour and intervals in the recognition of melody parts: evidence from cerebral asymmetries in musicians. *Neuropsychologia* 30 (3), 277–292.
- Peretz, I., Zatorre, R.J., 2005. Brain organization for music processing. *Annu. Rev. Psychol.* 56, 89–114.
- Platel, H., Price, C., Baron, J.C., Wise, R., Lambert, J., Frackowiak, R.S., Lechevalier, B., Eustache, F., 1997. The structural components of music perception. A functional anatomical study. *Brain* 120 (Pt 2), 229–243.
- Posner, 2004. *Cognitive Neuroscience of Attention*. Guilford Press, New York.
- Posner, Petersen, S.E., Fox, P.T., Raichle, M.E., 1988. Localization of cognitive operations in the human brain. *Science* 240 (4859), 1627–1631.
- Raizada, R.D.S., Tsao, F.M., Liu, H.M., Kuhl, P.K., 2010. Quantifying the adequacy of neural representations for a cross-language phonetic discrimination task: prediction of individual differences. *Cereb. Cortex* 20 (1), 1–12.
- Saarienen, J., Paavilainen, P., Schoger, E., Tervaniemi, M., Naatanen, R., 1992. Representation of abstract attributes of auditory stimuli in the human brain. *NeuroReport* 3 (12), 1149–1151.
- Schiavetto, A., Cortese, F., Alain, C., 1999. Global and local processing of musical sequences: an event-related brain potential study. *NeuroReport* 10 (12), 2467–2472.
- Schirmer, A., Kotz, S.A., 2006. Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends Cogn. Sci.* 10 (1), 24–30.
- Schon, D., Anton, J.L., Roth, M., Besson, M., 2002. An fMRI study of music sight-reading. *NeuroReport* 13 (17), 2285–2289.
- Schroeder, C.E., Foxe, J.J., 2002. The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. *Cogn. Brain Res.* 14 (1), 187–198.
- Shaywitz, S.E., Shaywitz, B.A., Fulbright, R.K., Skudlarski, P., Mencl, W.E., Constable, R.T., Pugh, K.R., Holahan, J.M., Marchione, K.E., Fletcher, J.M., Lyon, G.R., Gore, J.C., 2003. Neural systems for compensation and persistence: young adult outcome of childhood reading disability. *Biol. Psychiatry* 54 (1), 25–33.
- Stewart, L., Overath, T., Warren, J.D., Foxton, J.M., Griffiths, T.D., 2008. fMRI evidence for a cortical hierarchy of pitch pattern processing. *PLoS ONE* 3 (1), e1470.
- Stokes, M., Thompson, R., Nobre, A.C., Duncan, J., 2009. Shape-specific preparatory activity mediates attention to targets in human visual cortex. *Proc. Natl Acad. Sci. USA* 106 (46), 19569–19574.
- Temple, E., 2002. Brain mechanisms in normal and dyslexic readers. *Curr. Opin. Neurobiol.* 12 (2), 178–183.
- Tervaniemi, M., Maury, S., Naatanen, R., 1994. Neural representations of abstract stimulus features in the human brain as reflected by the mismatch negativity. *NeuroReport* 5 (7), 844–846.
- Trainor, L.J., Trehub, S.E., 1992. A comparison of infants' and adults' sensitivity to western musical structure. *J. Exp. Psychol. Hum. Percept. Perform.* 18 (2), 394–402.
- Trainor, L.J., Trehub, S.E., 1994. Key membership and implied harmony in Western tonal music: developmental perspectives. *Percept. Psychophys.* 56 (2), 125–132.
- Trainor, L.J., McDonald, K.L., Alain, C., 2002. Automatic and controlled processing of melodic contour and interval information measured by electrical brain activity. *J. Cogn. Neurosci.* 14 (3), 430–442.
- Walther, D.B., Caddigan, E., Fei-Fei, L., Beck, D.M., 2009. Natural scene categories revealed in distributed patterns of activity in the human brain. *J. Neurosci.* 29 (34), 10573–10581.
- Warren, J.D., Griffiths, T.D., 2003. Distinct mechanisms for processing spatial sequences and pitch sequences in the human auditory brain. *J. Neurosci.* 23 (13), 5799–5804.
- Warren, J.D., Uppenkamp, S., Patterson, R.D., Griffiths, T.D., 2003. Separating pitch chroma and pitch height in the human brain. *Proc. Natl Acad. Sci. USA* 100 (17), 10038–10042.
- Warrier, C.M., Zatorre, R.J., 2004. Right temporal cortex is critical for utilization of melodic contextual cues in a pitch constancy task. *Brain* 127 (Pt 7), 1616–1625.
- Zatorre, R.J., 1985. Discrimination and recognition of tonal melodies after unilateral cerebral excisions. *Neuropsychologia* 23 (1), 31–41.
- Zatorre, R.J., Evans, A.C., Meyer, E., 1994. Neural mechanisms underlying melodic perception and memory for pitch. *J. Neurosci.* 14 (4), 1908–1919.
- Zatorre, R.J., Halpern, A.R., Perry, D.W., Meyer, E., Evans, A.C., 1996. Hearing in the mind's ear: a PET investigation of musical imagery and perception. *J. Cogn. Neurosci.* 8 (1), 29–46.