AMERICAN PSYCHOLOGICAL ASSOCIATION ibnec Psychology & Neuroscience

ISSN: 1983-3288

2025, Vol. 18, No. 1, 40–59
https://doi.org/10.1037/pne0000354

Localizing Music's "Language of Emotion" in the Human Brain: A Functional MRI Study of Scale and Emotion Processing

Steven Brown, Matthew Berry, and Elizabeth Phillips
Department of Psychology, Neuroscience and Behaviour, McMaster University

Objective: The great mystery of music is the unique manner in which it is able to convey emotion. Its most domain-specific mechanism for doing so is tonality, most notably scale structure. Music's tonal structure creates a "language of emotion" whereby different scale types connote differences in emotional interpretation. In order to explore the neural basis of music's language of emotion, we carried out a functional magnetic resonance imaging study. ***Method:*** Trained musicians ($n = 31$) were tasked with discriminating the musical scale used in melodic samples, as well as the emotion conveyed by these samples, where the samples differed in the scale used (either major, minor, or chromatic). This was compared with a speech prosody condition in which participants had to discriminate the emotion conveyed in spoken utterances. ***Results:*** This comparison revealed the importance of regions that are little described in the neuromusic literature, namely, the lateral frontopolar cortex (Brodmann area 10/46) and the dorsal anterior cingulate cortex (area 32/8). ***Conclusions:*** The lateral frontopolar cortex and the dorsal anterior cingulate cortex contribute to the perception of emotional meaning in music, as conveyed through scale structure.

Public Significance Statement
The great mystery of music is the unique manner in which it is able to convey emotion. Its most unique mechanism for doing so is tonality, most notably scale structure. Music's tonal structure creates a "language of emotion" whereby different scale types connote differences in emotional interpretation. The results of our functional magnetic resonance imaging study demonstrate that brain areas such as the frontopolar cortex and anterior cingulate cortex contribute to the perception of emotional meaning in music.

Keywords: music, emotion, brain, tonality, musical scales

The topic of "music and emotion" covers two distinct—but oftentimes conflated—cognitive phenomena. One consists of the emotions that are *induced* in people when they listen to or perform music (Koelsch, 2018, 2020). The other consists of the emotions that are *conveyed* by music as a communicative device, in other words the emotions that are contained in the musical

---

Daniel Mograbi served as action editor.

The data sets supporting this article have been uploaded to the Dryad data repository at https://datadryad.org/stash/share/WpFFV57cnXYo3_rscgxVBGtGjPLwC9sI1gwkubr8BqQ. The authors have no conflicts of interest to disclose. This work was supported by a grant to Steven Brown from the Natural Sciences and Engineering Research Council of Canada (Grant RGPIN-2020-05718).

Steven Brown played a lead role in conceptualization, funding acquisition, investigation, project administration, resources, supervision, writing–original draft, and writing–review and editing and a supporting role in data curation, formal analysis, methodology, validation, and visualization. Matthew Berry played a lead role in data curation and formal analysis, a supporting role in investigation and writing–review and editing, and an equal role in methodology. Elizabeth Phillips played a supporting role in methodology and writing–review and editing.

Correspondence concerning this article should be addressed to Steven Brown, Department of Psychology, Neuroscience and Behaviour, McMaster University, 1280 Main Street West, Hamilton, ON L8S 4K1, Canada. Email: stebro@mcmaster.ca

object itself (Davies, 1994, 2001; Hatten, 2018; Kivy, 1980, 1990). Most of the neurocognitive literature on music and emotion has focused on the induction of emotions in response to music listening, such as the elicitation of aesthetic responses (Blood et al., 1999; Blood & Zatorre, 2001; Konečni, 2008) or the evocation of basic emotions like happiness and fear (Koelsch, 2020; Koelsch et al., 2013; Mitterschiffthaler et al., 2007). By contrast, the study of the emotions conveyed by music itself has received far less attention in the neuromusic literature, despite its central importance to music as a communication system. A primary function of music in human cultures is to enhance the emotional meaning of socially relevant information, most especially in the context of group rituals. The targets of such enhancement include verbal texts (Lomax, 1968; Sborgi Lawson, 2023), narrative scenarios (Cohen, 2013, 2015; Gorbman, 1987; Tan et al., 2017), and the ritual context itself (Arom & Khalfa, 1998; Merriam, 1964). In this regard, music operates similar to a semantic system, although one of *affective* semantics (Asano et al., 2022), rather than object-based semantics. Composers have to choose the musical-affective devices that are best suited for conveying the intended emotional meaning to an audience, whether that be in a song, dance, film, or ceremony.

Music has two systems for conveying emotion to listeners: prosody and tonality. Prosody refers to the expressive modulations of pitch height, melodic contour, tempo, and amplitude to communicate information about the intensity of the conveyed emotions (Clarke, 1988; Palmer & Hutchins, 2006). For example, music that is high-pitched, rising, fast, and loud is indicative of higher emotional intensity than music that is low-pitched, falling, slow, and soft. By all accounts, prosody in music shares most or all of its signaling features with prosody in speech. Numerous studies have shown that listeners attribute the same emotional interpretations to prosodic profiles, regardless of whether they are conveyed in music or speech (Juslin & Laukka, 2003). This makes vocal prosody a reasonable candidate for being an ancestral form of emotional communication that contributed jointly to music and speech during human evolution (S. Brown, 2000, 2017; Filippi, 2016; Fitch, 2013; Mithen, 2005; Ravignani & De Boer, 2021; Wallaschek, 1891), a coevolutionary hypothesis first proposed by Rousseau (1781/1998) in the 18th century. Peretz et al. (2015), in discussing the neural overlap between speech and music, considered the evolutionary possibility that “musicality recycles emotion circuits that have evolved for emotional vocalizations” (p. 6), an idea consistent with the coevolutionary account.

In contrast to this cross-modal—and most likely ancestral—nature of prosody, tonality is music’s most domain-specific and evolutionarily novel feature. By tonality we mean the organization of musical systems into scales, themselves comprised of collections of relatively discrete pitch classes and interval classes that are used recurrently in the generation of musical melodies and harmonies. A fundamental feature of tonality is that different scale types have different emotional-valence connotations. The best known example of this in Western music is the association of the major scale with positive emotional valence and the minor scale with negative valence (Huron, 2008; Parncutt, 2014). Such scale/emotion associations create a “language of emotion” that comprises music’s most domain-specific feature. This system is independent not only of prosody but also of other mechanisms of emotional communication, such as facial expression and body gesturing (Berry & Brown, 2022; Berry et al., 2022; de Gelder et al., 2015; Ekman & Friesen, 1974). In addition, scale/emotion associations are found not only in Western music but in the musical traditions of India, China, Japan, the Middle East, Persia, and Indonesia, among others (Malm, 1996). This is so despite the diversity of scale types found across these regions. Such associations make music into a powerful cultural device for communicating emotion, especially in ritual contexts. Music’s tonal mechanism for conveying emotional meanings is, at once, music’s defining cognitive feature and the essence of its mysterious allure.

A number of studies have made proposals for brain areas exhibiting neural specificity for music, highlighting regions such as the right posterior superior temporal gyrus (STG; Zatorre et al., 2002), the mid-STG bilaterally (Boebinger et al., 2021; Norman-Haignere et al., 2015), and the anterior STG bilaterally (Angulo-Perkins & Concha, 2019; Angulo-Perkins et al., 2014).

A recent meta-analysis attempted to look at the localization of tonality in the human brain. Asano et al. (2022) carried out an activation likelihood estimation (ALE) meta-analysis of 20 neuroimaging studies of tonal processing in music, most of them related to aspects of harmony processing (e.g., chord progressions, cadences). The results revealed a right-lateralized fronto-insular network, highlighting concordant activations across the 20 studies in (a) the right anterior insula and adjacent inferior frontal gyrus (IFG) pars orbitalis in Brodmann area (BA) 47, (b) the right BA 44 in the frontal operculum, and (c) the auditory association cortex of the anterior STG (BA 22) in the left hemisphere. A qualitative comparison of these results with Belyk and Brown's (2014) ALE meta-analysis of affective speech prosody—that is, the conveyance of emotion in spoken utterances separate from their linguistic content per se—revealed that virtually all of the peaks in the tonality analysis overlapped with those in the prosody analysis, bringing into question the music specificity of Asano et al.'s (2022) results. The prosody analysis demonstrated bilateral activations in these areas, whereas the tonality analysis tended to show right hemisphere areas alone, most especially in the anterior insula and IFG.

Neural overlap between musical tonality and affective speech prosody is also seen in experimental and lesion studies that have compared them directly. Merrill et al. (2012) carried out a comparative functional magnetic resonance imaging (fMRI) analysis of the prosodic aspect of speech and pitch processing in vocal music in a passive listening study. They observed a differential lateralization effect such that left BA 47 was associated with prosodic pitch processing in speech, whereas right BA 47 was associated with tonal pitch processing in music. An area that they found to be uniquely associated with music was the cortex of the intraparietal sulcus (IPS) bilaterally. Next, Sihvonen et al. (2022) carried out a comparative analysis of affective prosody and music processing in patients with either left hemisphere or right hemisphere strokes, using structural magnetic resonance imaging (MRI) to detect lesion locations. Similar lesion patterns for aprosodia and amusia were found in BA 47 in the right hemisphere and in the right striatum. In addition, an analysis of white matter tracts using diffusion tractography revealed common damage in the right inferior frontal occipital fasciculus, indicative of impaired processing along the ventral auditory pathway in both aprosodia and amusia.

The current functional MRI study examines the localization of scale/emotion processing in the brain by analyzing tasks involving musical scales and the emotions conveyed by them. Only a handful of previous neuroimaging studies have investigated scale or scale/emotion processing in music. Khalfa et al. (2005) had participants listen to Western classical musical samples (played with a piano timbre) that were either in the major or minor key. Their task was to judge the emotion conveyed in the music, along a continuum from sad to happy, as related to the scale used. The minor versus major contrast revealed activity in the left frontopolar (FP) cortex (BA 10), superior frontal gyrus (BA 9), and posterior cingulate cortex (BA 31). Green et al. (2008) had nonmusician participants listen to monophonic melodies in a major or minor key, or an ascending chromatic scale. Their task was to rate how much they liked each melody. Similar to the results of Khalfa et al. (2005), the minor versus major contrast in this study revealed activity in the left frontopolar cortex, as well as in the left parahippocampal gyrus and ventral anterior cingulate cortex (ACC) bilaterally. Tabei (2015) had music majors listen to musical melodies and judge the perceived emotion as being happy, neutral, or sad. A comparison of this to a passive listening condition revealed activity in the right IFG (BA 45) and left BA 47. Brattico et al. (2011) had participants make happy/sad emotion ratings on self-selected musical pieces that either contained or lacked lyrics. In looking at the main effect of emotion, they found that the sad versus happy contrast led to activity in the thalamus and caudate nucleus, whereas the happy versus sad contrast led to activity in the auditory cortex and a middle part of the insula. Bogert et al. (2016) had participants listen to musical samples from a movie soundtrack database that were meant to convey either happiness, sadness, or fear. In the "explicit" condition, participants had to rate the conveyed emotion of the sample, whereas in the "implicit" condition, they had to determine how many instruments they heard. The contrast of explicit versus implicit revealed activity in Lobule VI of the cerebellum bilaterally and the superior frontal gyrus (BA 9) bilaterally. In a very different approach to analyzing the neural basis of tonality, Janata et al. (2002) used a design in which the key of harmonic musical stimuli changed across the

12 major and 12 minor keys over the course of a scan. Brain areas that tracked changes in key included parts of the frontopolar cortex both ventral and dorsal to those described by Khalfa et al. (2005) and Green et al. (2008) in their minor versus major contrasts.

In order to examine the neural basis of music's "language of emotion" and to explore potential neural specificity for music in the human brain, we carried out a functional MRI study of tonal processing in music. Trained musicians listened to a series of piano melodies composed expressly for this study and had to perform two types of discrimination tasks based on tonal features of the sample: the scale used (major, minor, or chromatic) and the emotion conveyed (happy, calm, anxious, or sad). To the best of our knowledge, this is the first time that an explicit scale discrimination task has been carried out in a neuroimaging study. The samples controlled for prosodic features related to register, loudness, and tempo so that any perceived differences in emotional meaning were due to tonal features alone. In order to examine potential music specificity in the activation profile, not least in light of the overlap shown between tonal processing and speech prosody in the aforementioned studies, we also had participants listen to a series of semantically neutral spoken utterances and discriminate the affective prosody of the speaker, using the same four emotion categories as in the emotion task for music (happy, calm, anxious, or sad). Finally, we examined a musical task that exploits a prosodic parameter that is common to music and speech, namely, the contour of the melodies (rising, falling, or arched), which should be achievable without any explicit consideration of the sample's scale (Dowling, 1978). Based on the results of the tonality ALE meta-analysis and the aforementioned music/speech studies, we predicted that significant areas of overlap between musical tonality and affective speech prosody would occur in the anterior insula/IFG pars orbitalis (BA 47), frontal operculum, and anterior part of the superior temporal gyrus (aSTG), with an emphasis on the right hemisphere for the frontal areas. We also predicted that potential areas of music specificity for perceiving emotion in music might sit downstream of these areas, perhaps in the frontopolar cortex, as reported in previous studies of the perception of musical emotion and key changes.

Method

Participants

A gender-balanced group of 31 participants (16 females, 15 males), with a mean age of 29.5 ± 8.6 years, participated in the study after giving written informed consent (Hamilton Integrated Research Ethics Board, St. Joseph's Hospital). This sample size exceeds that of comparable fMRI studies of music and emotion (e.g., Bogert et al., 2016; Green et al., 2008; Khalfa et al., 2005; Seger et al., 2013). Each individual was without neurological or psychiatric illness. All participants except two (one female and one male) were right handed. All were either native English speakers or early arrivals in Canada. They were recruited by means of word of mouth, and were compensated monetarily for their participation. All participants were trained musicians, having begun musical training at a mean age of 8.0 ± 3.3 years old.

Stimuli

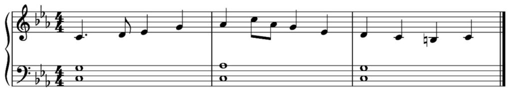
Two types of stimuli were generated for this study: musical melodies and spoken sentences. (a) *Melodies*. In order to generate 90 stimulus melodies—30 each in the major, minor, and chromatic scales—we used 10 base melodies in a major key for each of the three contour types (rising, falling, and arched), which were later transposed into the minor and chromatic scales (see below). Author E. P. composed all of the melodies in a major key, producing them using a piano timbre in MuseScore 3. The melodies contained, on average, 13.2 notes (*SD* 1.8, range = 10–17 notes) and had a mean duration of 5.6 s (*SD* 0.3 s, range = 5–6 s). All samples were within a 2.5-octave register from D#3 to A5, and had a mean melodic range of 16.4 semitones. Rhythms were balanced between quadruple (4/4) and triple (3/4) meters. Samples had no fewer than 2 durational values but no more than 3 (including dotted notes), with the basic duration value being a quarter note. The mean tempo was 120 bpm to the quarter note (range = 95–145 bpm).

In order to make the tonal features of the stimuli more apparent in the context of scanner noise, we added a small amount of harmonization involving 2-note chords spanning no more than a major 7th in size. There were one or two such chords per measure, resulting in three–six chords per sample

(M = 4 chords). A typical example of this is shown in Figure 1. In all cases, the harmonizations lacked the third degree of the scale so that the major/minor discrimination would only be discernable from the melodies and not from the harmonizations. Cadence types were not controlled for. In addition, interval categories were not controlled for. The arched melodies contained slightly larger melodic intervals than the ascending and descending melodies, although this should not have any impact on scale discrimination or even contour discrimination.

To create the minor-key version of the stimuli, the major-key versions were converted into the harmonic minor, containing a raised seventh in both ascent and descent. To create the chromatic (atonal) version of the stimuli, we made modifications to the major-key stimuli as follows. Ninety random number sequences were generated, comprised of +2, +1, 0, −1, and −2. Each melody was paired with one random sequence. Every second note in the melody, starting from the second pitch, was modified by the number of semitones indicated by the random sequence. In addition, one note of each chord, alternating between the upper and lower note, was modified by the number of semitones indicated. This typically resulted in seven pitch modifications in the melody and four pitch modifications in the harmonization. In order to preserve the original melodic contour during this process, we occasionally switched the contour-altering modification to the opposite polarity of the designated change (e.g., changed +2 to −2) or, if necessary, modified the size of the change by one semitone (e.g., +2 to +1). As a final step, we varied the keys of the 90 samples in order to eliminate the participants' employment of absolute pitch in making discriminations. Keys were assigned randomly within a range of ±2 semitones from the base key of C (256 Hz), in other words B♭, B, C, D♭, or D. By means of validation testing (n = 43 participants not taking part in the fMRI study, M_{age} = 18.9 years old), we picked the most unambiguously rated 60 stimuli from the 90 generated melodies. No musical stimulus was repeated across runs. In other words, each of the three music runs contained 20 unique melodic stimuli, which were presented in a fixed random order across participants.

Figure 1
An Overview of the Four Tasks Performed in This Study

MUSIC TASKS

SCALE	**Major, Minor, Chromatic**
EMOTION	**Happy, Calm, Anxious, Sad**
CONTOUR	**Rising, Falling, Arched**

SPEECH TASKS

"His glasses are on the floor. I see a rug on the floor."

PROSODY **Happy, Calm, Anxious, Sad**

Note. The top part of the figure shows a typical musical melody composed for the study (minor key with an arched melodic contour). In different conditions, the participants had to discriminate either the scale, the conveyed emotion, or the melodic contour. The bottom part shows a typical semantically neutral sentence used for the speech conditions. Participants had to discriminate the affective prosody of the sample. Note that the response options for the emotion and prosody conditions are the same.

(b) *Sentences*. Thirty semantically neutral base sentences with a mean number of 11.7 words and 14.2 syllables were obtained from Ben-David et al. (2011). In order for the stimuli to have a 5- to 6-s duration, the 30 stimulus sentences were combined into 15 two-sentence pairs. The syllable number for the sentences was approximately matched to the note number for the melodies (∼14 in both cases). Author M. B. (who has worked as a professional actor; male, 31 years old) spoke each of the 15 sentence stimuli in four different prosodies according to a simple circumplex organization for emotion: happy (positive valence, high intensity), calm (positive valence, low intensity), anxious (negative valence, high intensity), and sad (negative valence, low intensity). This resulted in a total of 60 speech stimuli, of which 20 were used in the fMRI experiment and the rest for training and other behavioral experiments not reported here. Recording was done using a SteelSeries Arctis 7 headset with a ClearCast bidirectional microphone (0.0–10 kHz). Each sentence was recorded as a.wav file using Audacity. Light editing was done to reduce background noise. If needed, a sound file was either lengthened or compressed slightly using Audacity to make the sample timing approximately 6 s.

Tasks

Participants underwent a 1-hr training session on the day prior to the scanning session in order to learn how to perform the tasks in a highly controlled manner in a supine position with a minimum amount of head, face, and body movement. The stimuli used for training were different from those used in the fMRI experiment. A recording of MRI scanner noise was played in the background so as to give participants a sense of the noise that would be present when listening to the music and speech stimuli in the scanner. During the actual fMRI scanning, participants performed the tasks, each one according to a blocked design of 16 s of cross-hair fixation alternating with 20 s of a discrimination task during a 6' run.

During each 10-s trial, a stimulus was presented for 6 s, followed by a 3-s response time in which either three or four buttons (depending on the task) of an MRI-compatible button box were used to register responses. This was followed by a 1-s "ready" screen to prepare the participant for the next stimulus. During task blocks, two stimuli were independently presented per 20 s block, hence making this a blocked design, rather than an event-related design. Because the stimuli were presented in a random order, the second stimulus of a task pair was completely independent of the first. Participants wore MRI-compatible headphones throughout the experiment. We calibrated the volume of the stimulus presentation before scanning began in order to ensure that participants could hear the stimuli well. Behavioral performance on the discrimination tasks (see the Results section) confirmed that stimulus presentation was effective.

Participants were instructed to focus their vision on a cross-hair throughout each run. During fixation blocks, the cross-hair was embedded in an array of meaningless X's in order to equate the visual stimulation to the response screens used in the task blocks. During the response periods of the four tasks, the cross-hair was replaced by text indicating the selection of possible responses and the buttons associated with them. During the fixation blocks of these scans, the participant was instructed to press randomly assigned buttons on two occasions so as to control for the motoric activations during the task blocks.

Participants performed six tasks in the scanner, one task per run. Only four are described here, as the other two conditions were part of a separate study. (a) *Scale discrimination*: Participants listened to a series of musical stimuli and had to determine if the scale type of each one was either major, minor, or chromatic by pressing the appropriate button on a button box. (b) *Emotion discrimination*: Participants listened to a series of musical stimuli and had to determine if the conveyed emotion of each one (not their own felt emotion) was either happy, calm, anxious, or sad by pressing the appropriate button on a button box. (c) *Contour discrimination*: Participants listened to a series of musical stimuli and had to determine if the melodic contour of each one was either rising, falling, or arched (i.e., rising then falling) by pressing the appropriate button on a button box. (d) *Prosody discrimination*: Participants listened to a series of speech stimuli and had to determine if the conveyed emotion of each one was happy, calm, anxious, or sad by pressing the appropriate button on a button box.

Image Acquisition and Data Analysis

MRI imaging was conducted on a 3T GE scanner at the Imaging Research Centre at St. Joseph's Hospital. Functional images sensitive to the blood-oxygen-level-dependent (BOLD) signal were collected with a gradient-echo echo planar imaging pulse sequence using standard parameters (repetition time = 2000 ms, echo time = 45 ms, flip angle = 90 degrees, 39 slices per volume, 4 mm slice thickness, no slice gap, matrix size = 64 × 64, field of view = 24 cm, voxel size = 3.75 mm × 3.75 mm × 4 mm), covering the whole brain. A total of 180 brain volumes were acquired over 6 min of scan time, corresponding with 10 alternations between 16 s epochs of fixation and 20 s epochs of task. Anatomical T1-weighted images were collected for each participant (3D-fast spoiled gradient-echo, inversion recovery-prepped, inversion time = 900 ms, echo time = 3.22 ms, flip angle = 9 degrees, receiver bandwidth = 31.25 kHz, number of excitations = 1, 164 slices, 1 mm slice thickness, no slice gap, matrix size = 512 × 512, field of view = 24 cm).

Functional image analysis was conducted using Brain Voyager QX (Version 22.2 Brain Innovation). Each functional scan was normalized to the Talairach template (Talairach & Tournoux, 1988). During the preprocessing stage, the scan series was realigned and motion corrected, 3D spatial smoothing was performed using a Gaussian filter with a full-width-at-half-maximum kernel size of 4 mm, and a temporal high-pass filter was applied at a frequency of 0.0078 Hz (or 1/128 s) using the general linear model-Fourier algorithm. The BOLD response for each task was modeled as the convolution of a 20 s boxcar with a synthetic hemodynamic response function composed of two gamma functions. The six head motion parameters were included as nuisance regressors in the analysis. In a first-level fixed-effects analysis, beta weights associated with the modeled hemodynamic responses were computed to fit the observed BOLD signal time course in each voxel for each participant using the general linear model. In a second-level group analysis, images for each task-versus-fixation contrast were brought forward into a random effects analysis. These analyses were corrected for multiple comparisons using the false discovery rate (FDR) at $q < 0.05$ ($k = 12$). Talairach coordinates for the activation peaks were extracted using NeuroElf (https://neuroelf.net).

Subtractions, Conjunctions, and Regions of Interests

The contrasts of interest used to evaluate music specificity in this study were the subtractions between the tonal tasks and speech prosody, hence scale versus prosody and emotion versus prosody. To look for shared activations between music and speech, we ran conjunction analyses: scale ∩ prosody and emotion ∩ prosody. In order to look at how key brain regions from the subtraction analyses responded to the three scale types (major, minor, and chromatic), we created 2.5-mm radius spherical regions of interest (ROIs) bilaterally. For the frontopolar cortex, two ROIs were derived from the emotion versus prosody subtraction ($q < 0.05$): −27, 50, 16 and 42, 50, 13. A third frontopolar ROI in the right hemisphere alone was based on the scale versus prosody subtraction ($q < 0.05$): 48, 41, 22. For the dorsal anterior cingulate cortex, bilateral ROIs were derived from the emotion versus prosody subtraction ($q < 0.05$): −6, 26, 40 and 3, 20, 40. Pairwise two-tailed *t* tests were carried out with Benjamini–Hochberg correction for multiple comparisons based on the three planned comparisons (major vs. minor vs. chromatic).

Results

The behavioral data revealed that participant accuracy on the four discrimination tasks in the scanner was 88.0% for scale, 92.3% for emotion, 90.5% for contour, and 94.4% for prosody across the 95% of responses for which button presses could be registered (the other 5% being missed responses). For the scale task, errors mainly resulted from a confusion between minor and major, and between chromatic and minor. For the emotion task, errors mainly resulted from a confusion of minor with the positive-valenced emotions. For the contour task, errors mainly resulted from a confusion between ascending and arched.

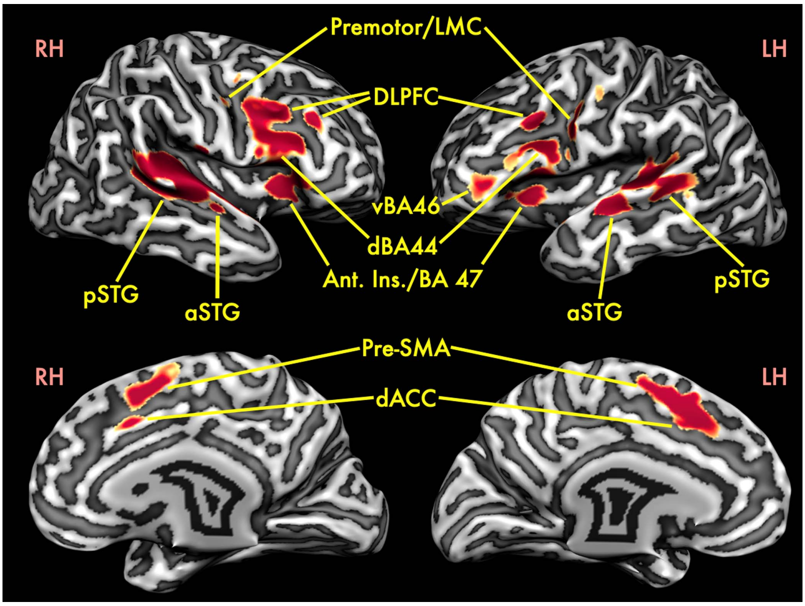
Figure 2 presents the group imaging results for the emotion versus fixation contrast registered onto the inflated brain of a single participant in order to demonstrate the basic brain network involved in performing the discrimination task. As mentioned in the Method section, button press was controlled for in the fixation epochs. The results of this contrast demonstrated a profile of bilateral brain areas collectively associated with the audiovocal network for music and speech processing (Belyk et al., 2016; S. Brown & Martinez, 2007). This included areas involved in low-level auditory perception (pSTG), high-level auditory processing (aSTG), and audiovocal interfacing (the anterior insula/BA 47, dorsal BA 44 [dBA44], presupplementary motor area [pre-SMA], premotor cortex, and the premotor part of the larynx motor cortex). It also included relatively domain-general areas involved in attentional/executive processing (dorsolateral prefrontal cortex, BA 46) that may underlie discrimination processing more generally.

In order to examine how much of this music network is shared with affective speech prosody, we ran conjunction analyses (Bonferroni $p < .05$) for both scale ∩ prosody and emotion ∩ prosody (Figure 3, with Talairach coordinates in Table 1). These analyses revealed that all of the general components of the network for emotion versus fixation were present in affective speech prosody as well, confirming the overlap between music and speech that has been reported in previous studies (see the Introduction).

The next step was to look for potential tonality-specific responses in light of this extensive overlap between music and prosody. To do so, we examined the subtractions of scale versus prosody and emotion versus prosody (FDR $q < 0.05$), as shown in Figure 4 and Table 2. Two areas were common across both subtractions: the lateral frontopolar cortex (BA 10 at the junction of ventral BA 46) and the dorsal part of the ACC (BA 32/8, just ventral to the pre-SMA). The emotion subtraction showed these areas bilaterally, while the scale subtraction emphasized the

Figure 2

Brain Areas for the Emotion Discrimination Task for Music



Note. Group results ($n = 31$) for emotion versus fixation are shown in both lateral (top) and medial (bottom) views in the right hemisphere (RH) and left hemisphere (LH), with Bonferroni ($p < .05$) correction ($k = 12$). Ant. Ins. = anterior insula; aSTG = anterior part of the superior temporal gyrus; BA = Brodmann area; dACC = dorsal part of the anterior cingulate cortex; dBA44 = dorsal part of BA 44; DLPFC = dorsolateral prefrontal cortex; LMC = larynx motor cortex (premotor part); pSTG = posterior part of the superior temporal gyrus; pre-SMA = presupplementary motor area; vBA46 = ventral part of BA 46.

right hemisphere alone for both areas. In addition, the scale location was directly dorsal to the main emotion peak ($z = 22$ compared to $z = 16$). In the right hemisphere, we thus distinguish "frontopolar (E)" from the emotion versus prosody subtraction and "frontopolar (S)" from the scale versus prosody subtraction. Next, the IPS was seen bilaterally in the scale versus prosody contrast, but not in the emotion versus prosody contrast. Surprisingly, neither the scale versus contour contrast nor the emotion versus contour contrast yielded significant peaks at corrected thresholds (not shown).

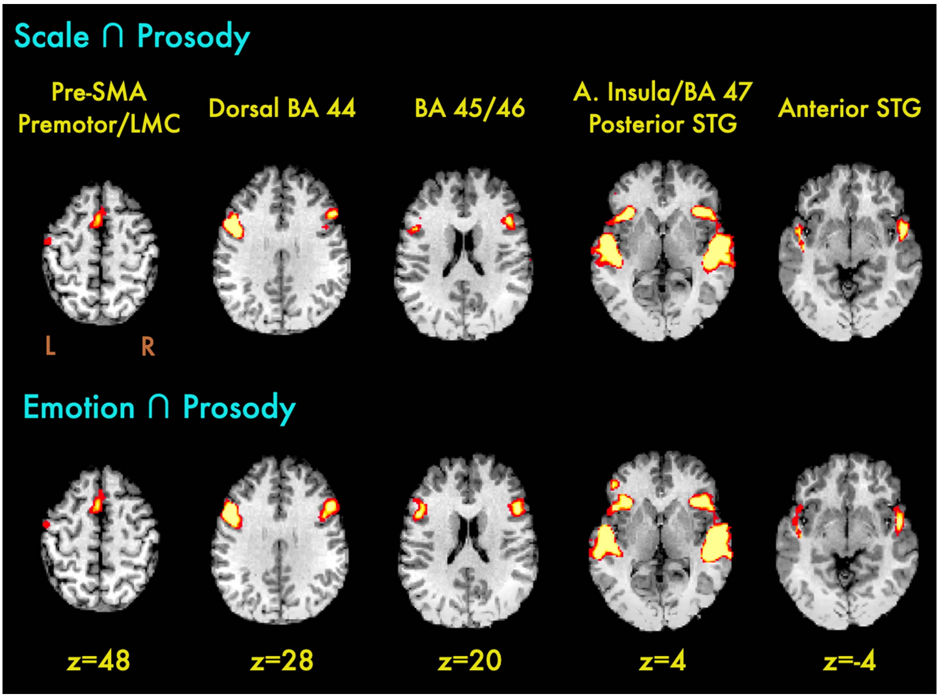
Finally, in order to examine how the frontopolar cortex and dorsal part of the anterior cingulate (dACC) responded to scale type in both the scale and emotion conditions, we obtained their peak coordinates from the emotion versus prosody subtraction (FDR $q < 0.05$)—as well as the right frontopolar peak from the scale versus emotion subtraction—and created ROIs in both hemispheres. The results are shown in Figure 5. For the scale condition (Panels A and B), the most reliable finding was that the major scale gave the lowest signal across the board. Major was significantly lower than both minor and chromatic (which did not differ between themselves) in the dACC, with a nonsignificant trend in the frontopolar cortex. The emotion condition (Panels C and D) presented a very different picture. The major scale now gave the highest signal across the board. It was significantly higher than chromatic and/or minor in three of the five ROIs, and trended in the same direction in the other two ROIs.

Discussion

In the neuroscientific study of music and emotion, conveyed emotion has received far less attention than felt emotion, despite the fact that music often functions to enhance the emotional meaning of socially relevant information. We used fMRI to identify brain areas that contribute to the domain-specific, and potentially human-specific,

Figure 3

Conjunction Analyses for Music and Affective Speech Prosody



Note. The results are shown with Bonferroni correction ($p < .05$). Below each row of slices is the Talairach z level. The left (L) side of the slice is the left side of the brain. A. Insula = anterior insula; BA = Brodmann area; LMC = larynx motor cortex (premotor part); pre-SMA = presupplementary motor area; STG = superior temporal gyrus.

communication of emotion via musical scales and the emotions connotatively associated with them, in other words to music's "language of emotion." In doing so, we carried out the first neuroimaging study of explicit scale discrimination, taking advantage of trained musicians as participants. The tonal tasks in this study were contrasted with the perception of emotion in spoken utterances via affective speech prosody. As predicted, there was extensive overlap between the neural systems for perceiving emotion in music and in speech, respectively. To the best of our knowledge, this is the first time that a conjunction analysis has been applied to active discrimination tasks for music and speech prosody, although Merrill et al.'s (2012) conjunction analysis of passive listening tasks observed overlap in the pSTG bilaterally. Our overlap included audiovocal areas typically associated with tonality (Asano et al., 2022), such as the anterior insula/BA 47, inferior frontal gyrus (BA 44), and auditory regions in the pSTG (BA 22) and aSTG (BA 22/38). Interestingly, the overlap in the anterior insula/BA 47 was found to be bilateral, despite expectations from the literature that it would mainly occur in the right hemisphere (Asano et al., 2022; Merrill et al., 2012; Sihvonen et al., 2022). The observation that the anterior insula/BA 47 is a point of strong overlap between music and speech prosody suggests that this area functions as a prosodic hub for music and speech, having little concern for either scaling or linguistic content. It is thus a good candidate for being a neural locus of the joint prosodic precursor of music and spoken language that several evolutionary models have proposed since the 18th century (S. Brown, 2000, 2017; Filippi, 2016; Fitch, 2013; Mithen, 2005; Ravignani & De Boer, 2021; Rousseau, 1781/1998; Wallaschek, 1891).

Beyond this overlap, the subtraction analyses between the music tasks and affective speech prosody revealed areas jointly important for scale

Table 1

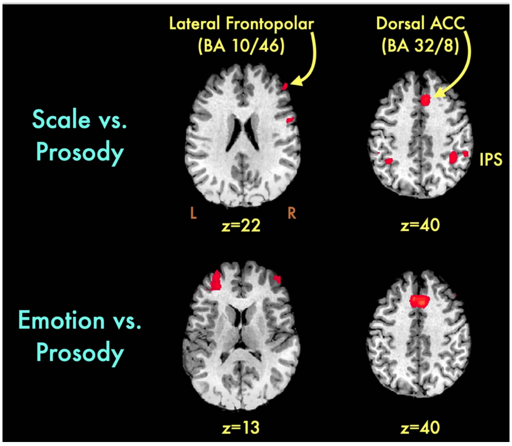
Talairach Coordinates for Scale ∩ Prosody and Emotion ∩ Prosody (Bonferroni $p < 0.05$; $k = 12$)

		Scale ∩ prosody					Emotion ∩ prosody				
Brain area	BA	*x*	*y*	*z*	Voxels	*t* max	*x*	*y*	*z*	Voxels	*t* max
Frontal and insula											
Right hemisphere											
Anterior insula	13	33	20	4	Cluster 1	10.24	33	23	7	Cluster 1	10.52
Precental gyrus	6	42	8	34	Cluster 1	9.50	42	8	34	Cluster 1	9.47
Dorsal IFG	45/9	48	20	28	Cluster 1	9.02	48	20	25	Cluster 1	10.47
	45	42	20	19	Cluster 1	8.99					
	44	45	8	22	Cluster 1	7.81					
Middle frontal gyrus	9	54	11	37	Cluster 1	6.97					
Left hemisphere											
Anterior insula	13	−33	17	4	Cluster 2	11.25	−30	20	4	Cluster 2	10.39
		−42	14	4	Cluster 2	9.03					
Dorsal IFG	44	−45	8	28	Cluster 2	12.64	−45	14	28	Cluster 2	12.49
	45	−48	26	19	Cluster 2	7.02	−51	26	19	Cluster 2	7.26
IFG							−42	17	10	Cluster 2	9.43
Premotor/LMC	6/4	−48	−7	46	Cluster 2	8.35	−48	−7	46	Cluster 2	7.54
ACC/pre-SMA	6	−3	14	46	87 (Cluster 3)	8.82	−3	14	46	91 (Cluster 3)	8.82
SMA	6	0	−1	58	Cluster 3	8.81	0	−1	58	Cluster 3	8.81
DLPFC	46	−45	38	7	15 (Cluster 4)	7.18	−45	35	7	Cluster 3	8.68
Dorsomedial prefrontal	8						0	26	49	Cluster 3	7.75
Temporal											
Right hemisphere											
Posterior STG	42	54	−16	7	858 (Cluster 1)	21.44	51	−16	7	952 (Cluster 1)	23.51
	22						54	−28	7	Cluster 1	18.07
Anterior STG	22	51	5	−5	Cluster 1	10.34					
Left hemisphere											
Posterior STG	42	−48	−22	7	812 (Cluster 2)	17.77	−48	−22	7	909 (Cluster 2)	17.64
	22	−51	−10	4	Cluster 2	14.66	−51	−13	4	Cluster 2	17.11
	41	−36	−31	13	Cluster 2	13.55					
	42	−60	−25	16	Cluster 2	10.70					
Anterior STG	22	−51	2	−2	Cluster 2	11.35	−51	11	1	Cluster 2	8.60
Cerebellum											
Lobule VI							−12	−76	−29	13 (Cluster 4)	7.84

Note. The maximal *t* value is shown in the *t* max column. BA = Brodmann area; IFG = inferior frontal gyrus; LMC = larynx motor cortex; ACC = anterior cingulate cortex; SMA = supplementary motor area; DLPFC = dorsolateral prefrontal cortex; STG = superior temporal gyrus.

Figure 4

Subtractions Between Music and Affective Speech Prosody



Note. The results are shown in the frontopolar cortex and dorsal ACC, with FDR correction ($q < 0.05$). The intraparietal sulcus (IPS) is only present in the scale versus prosody subtraction. Below each row of slices is the Talairach *z* level. The left (L) side of the slice is the left side of the brain. BA = Brodmann area; ACC = anterior cingulate cortex; FDR = false discovery rate.

discrimination and scale/emotion associations, namely, the lateral frontopolar cortex (BA 10/46) and the dorsal ACC (BA 32/8) bilaterally. The fact that these areas were seen comparably for the scale and emotion conditions suggests that the brain networks for scale discrimination and emotion discrimination are very similar. In fact, it is quite probable that when our musician participants performed the emotion task, they first engaged in a process of scale discrimination—either implicitly or explicitly—before selecting their choice of the emotion for a given sample. If this is indeed the case, then the emotion condition can be thought of as subsuming the scale condition, especially for this group of participants. The converse need not be the case, since nonmusicians demonstrate scale/emotion associations without possessing theoretical knowledge of scale categories (Castro & Lima, 2014; Dalla Bella et al., 2001; Hevner, 1935; Khalfa et al., 2005; Vieillard et al., 2008). Reciprocal subtraction between the scale and emotion conditions gave no signal at corrected thresholds (not shown). Despite the cognitive difference for musicians between using theory-driven knowledge to identify a musical scale and using more nontechnical knowledge to identify the emotion that is implicitly associated with the scale, the brain activations were very similar at the group level and even at the single-subject level. It is important to point out that we did not create the musical stimuli based on emotion categories but instead based on scale types. The discrimination of emotion was derived from connotative associations with the scale type.

The present study aimed at examining the neural encoding of the emotions conveyed by music. We can compare our results with Koelsch's (2020) ALE meta-analysis of the brain areas activated when people experience *felt* emotions in response to music listening. The networks between these two analyses are almost completely nonoverlapping, where the experience of felt emotions activates limbic areas such as the amygdala, hippocampus, striatum, orbitofrontal cortex, midcingulate cortex, ventral ACC, medial (but not lateral) frontopolar cortex, and various parts of the auditory cortex bilaterally. Hence, a comparison between conveyed emotion in the current fMRI study and felt emotion in Koelsch's meta-analysis demonstrates almost no similarity outside of auditory areas. This contrasts with the near-complete overlap between tonality and affective speech prosody seen in the present study (Figure 2), where both processes focus on perceived emotion. This is especially pertinent for the conjunction between emotion and prosody, where the response selections were identical.

The Frontopolar Cortex: Subdivisions, Functionality, and Phylogeny

Neither the lateral frontopolar cortex nor the dorsal ACC is well described in the neuromusic literature. In order to provide some context for these areas, we first consider Moayedi et al.'s (2015) resting-state functional connectivity analysis of the FP cortex. The FP cortex is divided into two functional subregions. The medial FP cortex (called FP2) is a component of the default mode network and shows connectivity with the posterior cingulate cortex. The lateral FP cortex (FP1) is associated with executive functioning and emotion. Importantly, its resting-state connectivity pattern shows a striking resemblance to the areas of activation seen in our subtraction analyses, namely, the anterior insula/BA 47 and the dorsal ACC. The only other area beyond these

Table 2

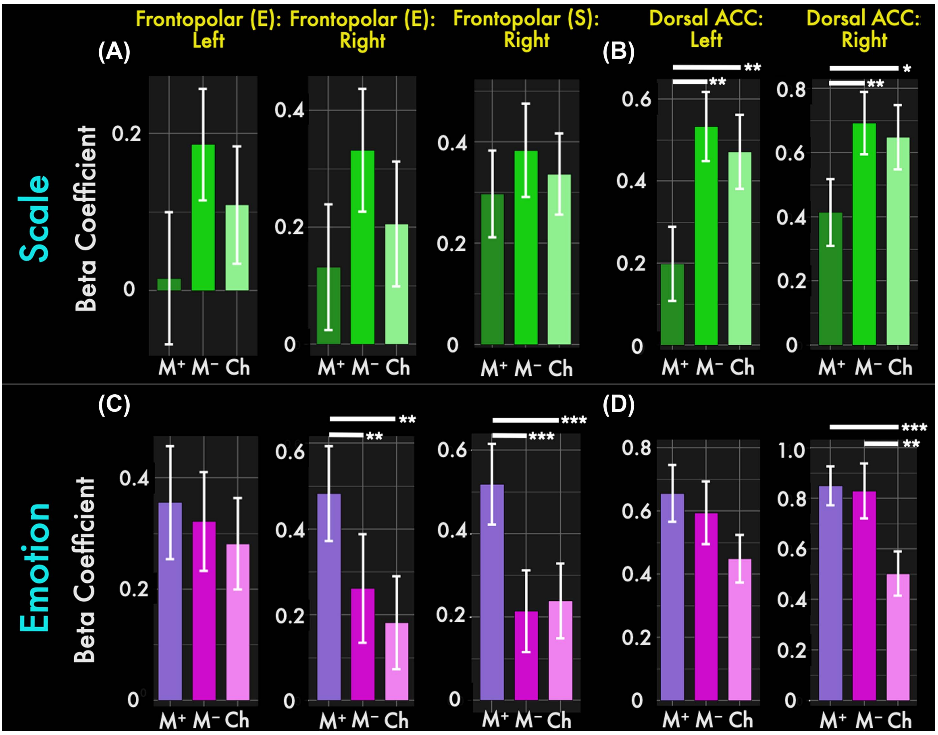
Talairach Coordinates for Scale Versus Prosody and Emotion Versus Prosody (FDR $q < 0.05$, $k = 12$)

		Scale versus prosody						Emotion versus prosody					
Brain area	BA	Hemis.	*x*	*y*	*z*	Voxels	*t* max	Hemis.	*x*	*y*	*z*	Voxels	*t* max
Dorsal ACC	32/8	RH	6	26	37	59 (Cluster 3)	5.35	RH	3	20	40	232 (Cluster 1)	5.92
		RH	15	29	28	Cluster 3	4.39	LH	−6	26	40	Cluster 1	5.87
								LH	−6	14	46	Cluster 1	5.26
								LH	−9	26	31	Cluster 1	5.22
								RH	9	11	46	Cluster 1	4.86
								RH	12	29	31	Cluster 1	4.23
Frontopolar cortex	10/46	RH	48	41	22	20 (Cluster 8)	4.55	LH	−27	50	16	69 (Cluster 2)	4.71
								LH	−24	41	16	Cluster 2	4.39
								LH	−36	38	10	Cluster 2	3.89
								RH	42	50	13	13 (Cluster 6)	4.11
Dorsolateral prefrontal	46/9							RH	39	29	28	69 (Cluster 3)	4.83
								RH	45	20	37	Cluster 3	3.93
								RH	39	20	25	Cluster 3	3.90
Intraparietal sulcus/IPL	40	LH	−33	−37	34	61 (Cluster 1)	5.09						
		LH	−36	−46	40	Cluster 1	4.65						
		LH	−27	−28	37	Cluster 1	4.31						
		LH	−45	−40	34	Cluster 1	3.98						
		RH	36	−43	43	61 (Cluster 2)	4.97						
		RH	54	−37	46	39 (Cluster 6)	4.48						
		RH	33	−64	43	14 (Cluster 10)	4.44						
SMA/pre-SMA	6	RH	27	−1	58	53 (Cluster 5)	4.91						
		LH	0	−10	64	20 (Cluster 9)	4.97						
Motor/premotor	6	RH	51	2	25	22 (Cluster 7)	4.80						
Cerebellum													
Lobule VI		LH	−27	−67	−20	57 (Cluster 4)	4.78	RH	30	−52	−32	33 (Cluster 4)	4.69
		LH	−18	−73	−11	Cluster 4	4.45	RH	36	−64	−23	Cluster 4	4.34
								LH	−30	−55	−29	16 (Cluster 5)	5.15

Note. The maximal *t* value is shown in the *t* max column. FDR = false discovery rate; BA = Brodmann area; Hemis. = hemisphere; ACC = anterior cingulate cortex; RH = right hemisphere; IPL = inferior parietal lobule; LH = left hemisphere; SMA = supplementary motor area.

Figure 5

Region-of-Interest Analysis for the Three Scale Types



Note. The scale condition is shown in Panels A–B, and the emotion condition is shown in Panels C–D. For the frontopolar ROIs, "Frontopolar (E)" refers to the left and right coordinates coming from the emotion versus prosody subtraction, whereas "Frontopolar (S)" refers to the right hemisphere coordinate coming from the scale versus prosody subtraction. Major gives the lowest signal in the scale condition, but the highest signal in the emotion condition. M+ = major scale; M− = minor scale; Ch = chromatic scale; ACC = anterior cingulate cortex; ROIs = regions of interest.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

two is the region of the intraparietal sulcus, which was seen in the scales versus prosody subtraction (see Figure 4) and which has been reported in a number of studies of active structural processing for music (R. M. Brown et al., 2013; Foster et al., 2013; Klein & Zatorre, 2011; Merrill et al., 2012), perhaps reflecting contour processing in our melodic tasks. Thus, all of the principal areas associated with tonal–emotional processing for music in the present study appear to be functionally interconnected. Gilbert et al. (2010), using a meta-analytic approach to connectivity, demonstrated that the rostral prefrontal cortex has connectivity with the anterior insula, dorsal ACC, and IPS, a profile similar to Moayedi et al. (2015). Whether these areas comprise a domain-general network for executive control or instead something more specific to emotion processing is in need of further analysis.

Interestingly, evolutionary neuroscientists have proposed that FP1 (the lateral FP cortex), but not FP2, might be a human-specific part of the brain. Neubert et al. (2014) carried out a comparative functional parcellation analysis of the brains of macaques and humans in the region of the ventral frontal cortex. They did this using diffusion-weighted tractography and an analysis of the connectivity fingerprint of each region with the rest of the brain. The functional connectivity pattern of FP2 was similar between humans and macaques, but that of FP1 was not. The authors argued that human FP1 "lacks simple correspondence with any [area] in macaque prefrontal cortex" (p. 708). These results were corroborated by Mars et al. (2016). The results suggest that FP1 may be a uniquely human brain area (Koch et al., 2018). According to Mansouri et al. (2017), the lateral FP subserves functions in humans that monkeys are less capable of performing. While it is too early to make strong associations between FP1 and tonality, it is important to point out that the use of musical scales and most especially the communication of emotion via scale/emotion associations appear to be human-specific functions. Therefore, it is not unreasonable to posit that their evolution might be subserved by new parts of the human brain, in connection with phylogenetically older areas like the anterior insula and IFG pars orbitalis.

Further insight into the role of the lateral FP cortex in emotion perception comes from an fMRI study of perceived emotion by Peelen et al. (2010). Participants observed actor-based portrayals of five basic emotions and had to rate the perceived emotional intensity of each one. In different blocks, they had to do this for stimuli presenting expression in either the voice, face, or body. The authors identified modality-independent neural representations of emotional state in the right FP cortex, although in a medial location to our right FP peaks for scale and emotion. A second cluster was found in the left posterior superior temporal sulcus. Peelen et al.'s (2010) task, by focusing on ratings of emotional intensity, is a bit closer to an emotional "prosody" task, whereas scale/emotion associations are more related to valence processing. However, their study shows that the lateral FP cortex has functions beyond its well-known associations with executive functioning, including the perception of emotion across the three principal modalities of expression.

Looking more specifically to music studies, the introduction mentioned the analyses of Khalfa et al. (2005) and Green et al. (2008) that observed activity in the left FP cortex in a location ventral to our area in their contrasts of minor versus major. In Green et al.'s (2008) study, the participants rated how much they liked the melodies they heard, and so any scale or emotion processing was implicit. By contrast, the study of Khalfa et al. (2005) had participants explicitly rate the perceived emotion in music samples. The results of these two studies suggest that the FP cortex is activated by both implicit and explicit tasks of perceived musical emotion, just as in our scale (implicit) and emotion (explicit) tasks. Bogert et al. (2016) found a peak in the left FP cortex lateral to the locations in the two previous studies for a task involving perceived musical emotion. What we observe is that all three of these studies show left hemisphere lateralization in the FP cortex. Might there be a lateralization effect here? In our own study, the scale recognition task only activated the right FP cortex, whereas the emotion discrimination task activated the FP cortex bilaterally. While it is far too early to claim that the right FP cortex is involved in scale processing and the left in scale/emotion associations—where the latter task would subsume the former in our cohort of trained musicians—this is an idea that can be explored in future studies of tonal processing.

Looking at other music studies not directly related to emotion or scales, Seger et al. (2013) observed activity in the left FP cortex when nonmusician participants listened to the cadences of musical pieces, compared to the initial segment of each piece. The results included other areas seen in our study, such as the anterior insula/IFG bilaterally, the IPS bilaterally, and an area directly anterior to our dACC. Lotze et al. (2003) found that the lateral FP cortex was more active bilaterally in amateur musicians than in professional musicians in tasks that involved both motor execution and musical imagery. In the execution task, the participants were supposed to perform the left-hand fingering of a Mozart violin concerto, whereas in the imagery condition, they were supposed to imagine doing so. Our own participant cohort was much closer to Lotze et al.'s (2003) amateur musicians than to their professional orchestra musicians. However, the results of their study demonstrate that activity in the FP cortex can be modulated by expertise. Finally, S. Brown and Martinez (2007) observed activity in the right lateral FP cortex during both melody discrimination for monophonic samples and harmony discrimination for harmonized melodies in a cohort of trained musicians similar to that of the present study.

The Dorsal Anterior Cingulate Cortex

The subtraction analyses revealed an effect of music at the intersection of the anterior cingulate cortex and dorsomedial prefrontal cortex at the junction of BA 32 and 8. Because the conjunction analysis demonstrated shared activity between music and speech prosody directly dorsal to the dACC in the pre-SMA region (see Figure 2), only the dACC showed a preference for music over speech prosody, whereas the region directly dorsal to this showed no such preference. As with the FP cortex, this region was not reported in Koelsch's (2020) ALE meta-analysis of brain areas associated with felt emotion.

Several neuromusic studies have reported activity in the dACC/dorsomedial prefrontal cortex across diverse tasks and participant cohorts, both musicians and nonmusicians. Bogert et al. (2016) reported activity in this area for a task involving musical emotion processing. However, they found it to be more active for implicit processing of emotion than explicit processing. Wallmark et al. (2018) reported a peak in this region during passive listening tasks when contrasting familiar and nonfamiliar music, but only when this contrast was regressed against trait empathy ("empathic concern"). Berns et al. (2010), in a study of adolescents (whose musicianship status was not reported), found that activation in the dACC correlated with participants' likability ratings for pop song samples. R. M. Brown et al. (2013), in a study using a repetition suppression paradigm in highly trained pianists, observed linear decreases in the dACC (what they called the pre-SMA) during a repetition condition for pitch sequences, but not one for timing sequences, when pianists had to engage in playback of heard melodies on a keyboard. As can be seen, there is little that unites these findings about the dACC across studies of music processing in both nonmusicians and trained musicians. At the present time, there is little to suggest that the dACC responds specifically to tonal aspects of music, compared to prosodic aspects, like melodic contour or other aspects of the structural processing of pitch in music.

The Effect of Scale Type

Using an ROI approach, we found that the major scale gave the lowest signal in the scale condition. Instead, the minor scale showed greater brain activity than the major scale (Figure 5), in support of the prior findings of Khalfa et al. (2005) and Green et al. (2008). In particular, we showed that the lateral FP cortex in both hemispheres responded more strongly to minor-key than major-key melodies in an explicit scale discrimination task, adding to this the observation that the chromatic scale showed intermediate responsiveness between minor and major. Whereas Green et al. (2008) found the chromatic scale to respond even more strongly than the minor in the FP cortex, we found the chromatic scale to be intermediate between major and minor. However, in three of the other ROIs that they analyzed, the chromatic scale was intermediate between major and minor. Our study revealed a similar scale effect in the dACC, which showed significantly higher activity for the minor and chromatic scales bilaterally.

Given that no ROI in the scale analysis showed significantly higher activity for the chromatic scale than the minor scale, we can most likely rule out the possibility that these brain areas respond more strongly to scales with greater numbers of scale tones than to those with fewer tones, in other words to scale complexity per se. In addition, since activity was higher for the chromatic scale than the major scale, these brain areas do not seem to have a preference for tonal music over atonal music, at least not in Western musician participants. However, it is not possible with the current data set to know if these brain areas respond more strongly to *all* nonmajor scales or just to the minor and chromatic scales. Further analyses using different scale types will be needed to address this issue. At the present time, we cannot rule out the possibility that the greater responsiveness to the minor and chromatic scales represents an "oddball" effect, compared to a default response to the major scale. The minor could simply be more difficult to detect. Behavioral analysis of response accuracy on the scale task in the scanner showed the following results: 96% major, 78% minor, and 88% chromatic. This might suggest that difficulty in discrimination could be one factor driving the activation level.

The ROI results for the emotion task painted a very different picture than those for the scale task, leading to a different cognitive interpretation. The major scale, instead of being the weakest scale, was now the strongest scale. The chromatic scale was significantly lower than the major or minor (or both) scales in all of the ROIs except for the right frontopolar ROIs. Given the fact that the participants consisted of trained musicians, it is very likely that they initially engaged in an implicit process of scale recognition when performing the emotion task—as was done explicitly in the scale task—and then made an emotion interpretation based on the recognized scale. Along these lines, the major and minor scales have equivalently strong associations with emotional-valence categories in Western culture, whereas the chromatic scale lacks such an association for most people, even in trained musicians. Hence, the comparable activity for the major and minor scales for the emotion task in three out of the five ROIs might stem from the similar degree of associability of these two scales to particular basic-emotion categories in Western musicians.

Limitations

Perhaps the biggest limitation of the present study is that we used trained musicians as the participants. This was done because nonmusicians would not be able to perform an explicit scale recognition task in a competent manner, although they should be able to perform the emotion task, since scale/emotion associations are learned implicitly through cultural exposure. The emotion task could thus be carried out in future studies with nonmusicians to see if the effects observed in the FP cortex and dACC are replicated in people with no explicit knowledge about scale categories. Note that the study of Khalfa et al. (2005) does not mention anything about the musicianship status of its participants and so may have already carried out such a study in nonmusicians.

Another limitation of the study is that we placed the emphasis on monophonic melodies, introducing only a small amount of harmonization. Because the previous studies on scales have used chordal samples, we have been able to replicate their findings without employing strong harmonization by simply using 2-note chords that lacked the third degree of the scale, thus forcing participants to rely on the melodic sequence to identify the scale type, especially for the major/minor distinction. However, we would like to know more about how harmony (or musical texture more generally) affects activity in areas like the FP cortex and dACC, especially given the fact that these areas respond differentially to scale types.

Given the divergent results between the scale and emotion tasks with regard to scale type (see Figure 5), it will be important to carry out studies with many more types of scales, most especially using participants from different cultures. For example, India is perhaps the most scale-rich country in the world, where theoretical analyses provide strong associations between the "raga" scales of the music and particular emotions (among other types of associations with the scales; Massey & Massey, 1993). Hence, not only would we want to know whether Indian participants respond the same way as our North American participants to the major and minor scales, but how they respond to the ragas that they have great exposure to, but which North American people have little to no exposure to. Hence, there is a great need for a cultural neuroscience approach (Chiao, 2009) to the study of musical scales, although there has been very little neuroimaging work on the perception of musical scales even in Western individuals.

Conclusions

The present study explored the neural correlates of music's language of emotion as conveyed through scale structure, seeking to identify brain areas distinct from the prosodic mechanisms shared with speech, as well as from nonacoustic mechanisms of emotion perception, such as via facial expression and body gesturing. Areas showing relatively stronger activity for both scale discrimination and the emotional interpretation of scales than for affective speech prosody included the lateral frontopolar cortex and the dorsal ACC bilaterally. These areas are separate from those that process the induction of felt emotions when people listen to music. These areas are not well described in the neuromusic literature. However, they are interconnected with one another in previous resting-state connectivity and meta-analytic connectivity studies and may thus comprise a functional network for perceiving conveyed emotions. The phylogenetic observation that the lateral frontopolar cortex (FP1) is potentially a human-specific area might provide some insight into the evolutionary novelty and domain specificity of musical scales and scale/emotion associations in the context of both human cognition and human evolution. Given the centrality of conveyed emotion to the social functions of music, it is surprising that musical scales have been so understudied in the neuromusic literature. In addition, scale structure is one of the most diverse features of musical systems cross-culturally. More attention needs to be directed at understanding music's language of emotion and how this novel communicative function is represented in the human brain and across human populations.

References

Angulo-Perkins, A., Aubé, W., Peretz, I., Barrios, F. A., Armony, J. L., & Concha, L. (2014). Music listening engages specific cortical regions within the temporal lobes: Differences between musicians and non-musicians. *Cortex*, *59*, 126–137. https://doi.org/10.1016/j.cortex.2014.07.013

Angulo-Perkins, A., & Concha, L. (2019). Discerning the functional networks behind processing of music and speech through human vocalizations. *PLOS ONE*, *14*(10), Article e0222796. https://doi.org/10.1371/journal.pone.0222796

Arom, S., & Khalfa, J. (1998). Une raison en acte: Pensée formelle et systématique musicale dans les sociétés de tradition orale [A reason in action: Formal thought and musical systematics in societies of oral tradition]. *Revue de Musicologie*, *84*(1), 5–17. https://doi.org/10.2307/947326

Asano, R., Lo, V., & Brown, S. (2022). The neural basis of tonal processing in music: An ALE meta-analysis. *Music & Science*, *5*, 1–15. https://doi.org/10.1177/20592043221109958

Belyk, M., & Brown, S. (2014). Perception of affective and linguistic prosody: An ALE meta-analysis of neuroimaging studies. *Social Cognitive and Affective Neuroscience*, *9*(9), 1395–1403. https://doi.org/10.1093/scan/nst124

Belyk, M., Pfordresher, P. Q., Liotti, M., & Brown, S. (2016). The neural basis of vocal pitch imitation in humans. *Journal of Cognitive Neuroscience*, *28*(4), 621–635. https://doi.org/10.1162/jocn_a_00914

Ben-David, B. M., Van Lieshout, P. H., & Leszcz, T. (2011). A resource of validated affective and neutral sentences to assess identification of emotion in spoken language after a brain injury. *Brain Injury*, *25*(2), 206–220. https://doi.org/10.3109/02699052.2010.536197

Berns, G. S., Capra, C. M., Moore, S., & Noussair, C. (2010). Neural mechanisms of the influence of popularity on adolescent ratings of music. *NeuroImage*, *49*, 2687–2696. https://doi.org/10.1016/j.neuroimage.2009.10.070

Berry, M., & Brown, S. (2022). The dynamic mask: Facial correlates of character portrayal in professional actors. *Quarterly Journal of Experimental Psychology*, *75*(5), 936–953. https://doi.org/10.1177/17470218211047935

Berry, M., Lewin, S., & Brown, S. (2022). Correlated expression of the body, face, and voice during character portrayal in actors. *Scientific Reports*, *12*(1), Article 8253. https://doi.org/10.1038/s41598-022-12184-7

Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(20), 11818–11823. https://doi.org/10.1073/pnas.191355898

Blood, A. J., Zatorre, R. J., Bermudez, P., & Evans, A. C. (1999). Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nature Neuroscience*, *2*(4), 382–387. https://doi.org/10.1038/7299

Boebinger, D., Norman-Haignere, S. V., McDermott, J. H., & Kanwisher, N. (2021). Music-selective neural populations arise without musical training. *Journal of Neurophysiology*, *125*(6), 2237–2263. https://doi.org/10.1152/jn.00588.2020

Bogert, B., Numminen-Kontti, T., Gold, B., Sams, M., Numminen, J., Burunat, I., Lampinen, J., & Brattico, E. (2016). Hidden sources of joy, fear, and sadness: Explicit versus implicit neural processing of musical emotions. *Neuropsychologia*, *89*, 393–402. https://doi.org/10.1016/j.neuropsychologia.2016.07.005

Brattico, E., Alluri, V., Bogert, B., Jacobsen, T., Vartiainen, N., Nieminen, S., & Tervaniemi, M. (2011). A functional MRI study of happy and sad emotions in music with and without lyrics. *Frontiers in Psychology*, *2*, Article 308. https://doi.org/10.3389/fpsyg.2011.00308

Brown, R. M., Chen, J. L., Hollinger, A., Penhune, V. B., Palmer, C., & Zatorre, R. J. (2013). Repetition suppression in auditory–motor regions to pitch and temporal structure in music. *Journal of Cognitive Neuroscience*, *25*(2), 313–328. https://doi.org/10.1162/jocn_a_00322

Brown, S. (2000). The "musilanguage" model of music evolution. In N. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 271–300). MIT Press.

Brown, S. (2017). A joint prosodic origin of language and music. *Frontiers in Psychology*, *8*, Article 1894. https://doi.org/10.3389/fpsyg.2017.01894

Brown, S., & Martinez, M. J. (2007). Activation of premotor vocal areas during musical discrimination. *Brain and Cognition*, *63*(1), 59–69. https://doi.org/10.1016/j.bandc.2006.08.006

Castro, S. L., & Lima, C. F. (2014). Age and musical expertise influence emotion recognition in music. *Music Perception*, *32*(2), 125–142. https://doi.org/10.1525/mp.2014.32.2.125

Chiao, J. Y. (2009). Cultural neuroscience: A once and future discipline. *Progress in Brain Research*, *178*, 287–304. https://doi.org/10.1016/S0079-6123(09)17821-4

Clarke, E. (1988). Generative principles in music performance. In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition* (pp. 1–26). Oxford University Press.

Cohen, A. J. (2013). Film music. In S.-L. Tan, A. J. Cohen, S. D. Lipscomb, & R. A. Kendal (Eds.), *The psychology of music in multimedia* (pp. 17–47).

Oxford University Press. https://doi.org/10.1093/acprof:Oso/9780199608157.003.0002

Cohen, A. J. (2015). Congruence-association model and experiments in film music: Toward interdisciplinary collaboration. *Music and the Moving Image*, *8*(2), 5–24. https://doi.org/10.5406/musimoviimag.8.2.0005

Dalla Bella, S., Peretz, I., Rousseau, L., & Gosselin, N. (2001). A developmental study of the affective value of tempo and mode in music. *Cognition*, *80*(3), B1–B10. https://doi.org/10.1016/S0010-0277(00)00136-0

Davies, S. (1994). *Musical meaning and expression*. Cornell University Press. https://doi.org/10.7591/9781501733987

Davies, S. (2001). Philosophical perspectives on music's expressiveness. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 23–44). Oxford University Press. https://doi.org/10.1093/oso/9780192631886.003.0002

de Gelder, B., de Borst, A. W., & Watson, R. (2015). The perception of emotion in body expressions. *Wiley Interdisciplinary Reviews: Cognitive Science*, *6*(2), 149–158. https://doi.org/10.1002/wcs.1335

Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review*, *85*(4), 341–354. https://doi.org/10.1037/0033-295X.85.4.341

Ekman, P., & Friesen, W. V. (1974). Detecting deception from the body or face. *Journal of Personality and Social Psychology*, *29*(3), 288–298. https://doi.org/10.1037/h0036006

Filippi, P. (2016). Emotional and interactional prosody across animal communication systems: A comparative approach to the emergence of language. *Frontiers in Psychology*, *7*, Article 1393. https://doi.org/10.3389/fpsyg.2016.01393

Fitch, W. T. (2013). Musical protolanguage: Darwin's theory of language evolution revisited. In J. J. Bolhuis & M. Everaert (Eds.), *Birdsong, speech, and language: Exploring the evolution of mind and brain* (pp. 489–504). MIT Press. https://doi.org/10.7551/mitpress/9322.003.0032

Foster, N. E. V., Halpern, A. R., & Zatorre, R. J. (2013). Common parietal activation in musical mental transformations across pitch and time. *NeuroImage*, *75*, 27–35. https://doi.org/10.1016/j.neuroimage.2013.02.044

Gilbert, S. J., Gonen-Yaacovi, G., Benoit, R. G., Volle, E., & Burgess, P. W. (2010). Distinct functional connectivity associated with lateral versus medial rostral prefrontal cortex: A meta-analysis. *NeuroImage*, *53*(4), 1359–1367. https://doi.org/10.1016/j.neuroimage.2010.07.032

Gorbman, C. (1987). *Unheard melodies: Narrative film music*. Indiana University Press.

Green, A. C., Baerentsen, K. B., Stødkilde-Jørgensen, H., Wallentin, M., Roepstorff, A., & Vuust, P. (2008). Music in minor activates limbic structures: A relationship with dissonance? *NeuroReport*, *19*(7), 711–715. https://doi.org/10.1097/WNR.0b013e3282fd0dd8

Hatten, R. S. (2018). *A theory of virtual agency for Western art music*. Indiana University Press. https://doi.org/10.2307/j.ctv512trj

Hevner, K. (1935). The affective character of the major and minor modes in music. *The American Journal of Psychology*, *47*(1), 103–118. https://doi.org/10.2307/1416710

Huron, D. (2008). A comparison of average pitch height and interval size in major- and minor-key themes: Evidence consistent with affect-related pitch prosody. *Empirical Musicology Review*, *3*, 59–63. https://doi.org/10.18061/1811/31940

Janata, P., Birk, J. L., Van Horn, J. D., Leman, M., Tillmann, B., & Bharucha, J. J. (2002). The cortical topography of tonal structures underlying Western music. *Science*, *298*(5601), 2167–2170. https://doi.org/10.1126/science.1076262

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, *129*(5), 770–814. https://doi.org/10.1037/0033-2909.129.5.770

Khalfa, S., Schon, D., Anton, J. L., & Liégeois-Chauvel, C. (2005). Brain regions involved in the recognition of happiness and sadness in music. *NeuroReport*, *16*(18), 1981–1984. https://doi.org/10.1097/00001756-200512190-00002

Kivy, P. (1980). *The corded shell: Reflections on musical experience*. Princeton University Press.

Kivy, P. (1990). *Music alone: Philosophical reflections on the purely musical experience*. Cornell University Press.

Klein, M. E., & Zatorre, R. J. (2011). A role for the right superior temporal sulcus in categorical perception of musical chords. *Neuropsychologia*, *49*(5), 878–887. https://doi.org/10.1016/j.neuropsychologia.2011.01.008

Koch, S. B. J., Mars, R. B., Toni, I., & Roelofs, K. (2018). Emotional control, reappraised. *Neuroscience and Biobehavioral Reviews*, *95*, 528–534. https://doi.org/10.1016/j.neubiorev.2018.11.003

Koelsch, S. (2018). Investigating the neural encoding of emotion with music. *Neuron*, *98*(6), 1075–1079. https://doi.org/10.1016/j.neuron.2018.04.029

Koelsch, S. (2020). A coordinate-based meta-analysis of music-evoked emotions. *NeuroImage*, *223*, Article 117350. https://doi.org/10.1016/j.neuroimage.2020.117350

Koelsch, S., Skouras, S., Fritz, T., Herrera, P., Bonhage, C., Küssner, M. B., & Jacobs, A. M. (2013). The roles of superficial amygdala and auditory cortex in music-evoked fear and joy. *NeuroImage*, *81*, 49–60. https://doi.org/10.1016/j.neuroimage.2013.05.008

Konečni, V. J. (2008). Does music induce emotion? A theoretical and methodological analysis. *Psychology of Aesthetics, Creativity, and the Arts*, *2*(2), 115–129. https://doi.org/10.1037/1931-3896.2.2.115

Lomax, A. (1968). *Folk song style and culture*. American Association for the Advancement of Science.

Lotze, M., Scheler, G., Tan, H. R. M., Braun, C., & Birbaumer, N. (2003). The musician's brain: Functional imaging of amateurs and professionals during performance and imagery. *NeuroImage*, *20*(3), 1817–1829. https://doi.org/10.1016/j.neuroimage.2003.07.018

Malm, W. P. (1996). *Music cultures of the Pacific, the Near East and Asia* (3rd ed.). Prentice Hall.

Mansouri, F. A., Koechlin, E., Rosa, M. G. P., & Buckley, M. J. (2017). Managing competing goals—A key role for the frontopolar cortex. *Nature Reviews Neuroscience*, *18*(11), 645–657. https://doi.org/10.1038/nrn.2017.111

Mars, R. B., Verhagen, L., Gladwin, T. E., Neubert, F. X., Sallet, J., & Rushworth, M. F. S. (2016). Comparing brains by matching connectivity profiles. *Neuroscience and Biobehavioral Reviews*, *60*, 90–97. https://doi.org/10.1016/j.neubiorev.2015.10.008

Massey, R., & Massey, J. (1993). *The music of India*. Kahn & Averill.

Merriam, A. P. (1964). *The anthropology of music*. Northwestern University Press.

Merrill, J., Sammler, D., Bangert, M., Goldhahn, D., Lohmann, G., Turner, R., & Friederici, A. D. (2012). Perception of words and pitch patterns in song and speech. *Frontiers in Psychology*, *3*, Article 76. https://doi.org/10.3389/fpsyg.2012.00076

Mithen, S. (2005). *The singing Neanderthals: The origins of music, language, mind and body*. Weidenfeld & Nicolson.

Mitterschiffthaler, M. T., Fu, C. H. Y., Dalton, J. A., Andrew, C. M., & Williams, S. C. R. (2007). A functional MRI study of happy and sad affective states induced by classical music. *Human Brain Mapping*, *28*(11), 1150–1162. https://doi.org/10.1002/hbm.20337

Moayedi, M., Salomons, T. V., Dunlop, K. A. M., Downar, J., & Davis, K. D. (2015). Connectivity-based parcellation of the human frontal polar cortex. *Brain Structure & Function*, *220*(5), 2603–2616. https://doi.org/10.1007/s00429-014-0809-6

Neubert, F. X., Mars, R. B., Thomas, A. G., Sallet, J., & Rushworth, M. F. S. (2014). Comparison of human ventral frontal cortex areas for cognitive control and language with areas in monkey frontal cortex. *Neuron*, *81*(3), 700–713. https://doi.org/10.1016/j.neuron.2013.11.012

Norman-Haignere, S., Kanwisher, N. G., & McDermott, J. H. (2015). Distinct cortical pathways for music and speech revealed by hypothesis-free voxel decomposition. *Neuron*, *88*(6), 1281–1296. https://doi.org/10.1016/j.neuron.2015.11.035

Palmer, C., & Hutchins, S. (2006). What is musical prosody? In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 245–278). Elsevier.

Parncutt, R. (2014). The emotional connotations of major versus minor tonality: One or more origins? *Musicae Scientiae*, *18*(3), 324–353. https://doi.org/10.1177/1029864914542842

Peelen, M. V., Atkinson, A. P., & Vuilleumier, P. (2010). Supramodal representations of perceived emotions in the human brain. *The Journal of Neuroscience*, *30*(30), 10127–10134. https://doi.org/10.1523/JNEUROSCI.2161-10.2010

Peretz, I., Vuvan, D., Lagrois, M.-E., & Armony, J. L. (2015). Neural overlap in processing music and speech. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1664), Article 20140090. https://doi.org/10.1098/rstb.2014.0090

Ravignani, A., & De Boer, B. (2021). Joint origins of speech and music: Testing evolutionary hypotheses on modern humans. *Semiotica*, *2021*(239), 169–176. https://doi.org/10.1515/sem-2019-0048

Rousseau, J.-J. (1998). *Essay on the origin of languages* (J. T. Scott, Ed. & Trans.). University Press of New England. (Original work published 1781)

Sborgi Lawson, F. R. (2023). Mousike or music? Using analysis to explore shifts in musical attention. *Ethnomusicology Forum*, *32*(1), 120–142. https://doi.org/10.1080/17411912.2023.2168286

Seger, C. A., Spiering, B. J., Sares, A. G., Quraini, S. I., Alpeter, C., David, J., & Thaut, M. H. (2013). Corticostriatal contributions to musical expectancy perception. *Journal of Cognitive Neuroscience*, *25*(7), 1062–1077. https://doi.org/10.1162/jocn_a_00371

Sihvonen, A. J., Sammler, D., Ripollés, P., Leo, V., Rodríguez-Fornells, A., Soinila, S., & Särkämö, T. (2022). Right ventral stream damage underlies both poststroke aprosodia and amusia. *European Journal of Neurology*, *29*(3), 873–882. https://doi.org/10.1111/ene.15148

Tabei, K. (2015). Inferior frontal gyrus activation underlies the perception of emotions, while precuneus activation underlies the feeling of emotions during music listening. *Behavioural Neurology*, *2015*(1), Article 529043. https://doi.org/10.1155/2015/529043

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxix atlas of the human brain. 3-dimensional proportional system: An approach to cerebral imaging*. Gerg Thieme Verlag.

Tan, S. L., Spackman, M. P., & Wakefield, E. M. (2017). The effects of diegetic and nondiegetic music on viewers' interpretations of a film scene.

Music Perception, *34*(5), 605–623. https://doi.org/10.1525/mp.2017.34.5.605

Vieillard, S., Peretz, I., Gosselin, N., Khalfa, S., Gagnon, L., & Bouchard, B. (2008). Happy, sad, scary and peaceful musical excerpts for research on emotions. *Cognition and Emotion*, *22*(4), 720–752. https://doi.org/10.1080/02699930701503567

Wallaschek, R. (1891). On the origins of music. *Mind*, *16*(63), 375–386. https://doi.org/10.1093/mind/os-XVI.63.375

Wallmark, Z., Deblieck, C., & Iacoboni, M. (2018). Neurophysiological effects of trait empathy in music listening. *Frontiers in Behavioral Neuroscience*, *12*, Article 66. https://doi.org/10.3389/fnbeh.2018.00066

Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences*, *6613*, 37–46. https://doi.org/10.1016/S1364-6613(00)01816-7

Received May 14, 2024
Revision received September 15, 2024
Accepted September 16, 2024 ■