

The Neural Basis of Tonal Processing in Music: An ALE Meta-Analysis

Music & Science
Volume 5: 1–15
© The Author(s) 2022
DOI: 10.1177/20592043221109958
journals.sagepub.com/home/mns

Rie Asano¹ , Vivian Lo² and Steven Brown²

Abstract

Music is used as an important medium for communication in human societies, often times to enhance the emotional meaning of narrative scenarios and ritual events. Music has a number of domain-specific tonal devices for doing this, spanning from scale structure to harmonic progressions and beyond. In order to explore the neural basis of tonal processing in music, we carried out an activation likelihood estimation (ALE) meta-analysis of 20 published functional magnetic resonance imaging studies of tonal cognition, with an emphasis on harmony processing. The most concordant areas of activation across these studies occurred at the junction of the inferior frontal gyrus, anterior insula, and orbitofrontal cortex in Brodmann areas 47 and 13 in the right hemisphere. This region is associated not only with emotion in general, but with the conveyance of affective meanings during communication processes, including speech prosody and music.

Keywords

Chord, fMRI, harmony, melody, music, scale, speech prosody, tonality

Submission date: 19 November 2021; Acceptance date: 11 June 2022

Introduction

Tonality is one of the most domain-specific features of music. At its most basic level, tonality involves the use of sets of discrete pitches for the purpose of creating interval classes and musical melodies (Sachs, 1943). Such pitch sets can be formalized into scale types, modes, and keys that are organized according to tonal hierarchies in which different scale positions have different functional roles in the generation of melodies (Krumhansl, 1990; Krumhansl & Kessler, 1982; Lerdahl, 2001). The notion of a tonal hierarchy applies not only to monophonic melodies but to the textural level of music, especially to homophonic texture, in which different chord classes have distinct functional roles in the generation of harmonic progressions (Krumhansl, 1990; Schoenberg, 1922). It is this series of melodic and harmonic tonal devices related to hierarchical organization that theorists associate with the syntax of music (Lerdahl, 2013; Patel, 2008). Figure 1 provides an overview of some of the most important tonality-related concepts in music theory.

In music, tonality is intimately associated with the conveyance of affect. For example, tonality, especially the tonal hierarchy, is a driving force for the encoding of tension-relaxation patterns in music (Jackendoff & Lerdahl, 2006). The results

of both theoretical and empirical research suggest that tonal-expectancy violations, as well as transitions from stable to less stable pitches, can cause tension, whereas fulfilled expectancies and resolutions towards a stable pitch can lead to relaxation (Krumhansl, 2002; Steinbeis et al., 2006; Steinbeis & Koelsch, 2008). Another central 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 mode with positive emotional valence and the minor mode with negative valence (Huron, 2008; Parncutt, 2014). This valence-based distinction tends to be categorical, in contrast to the graded manner in which emotional intensity is encoded in music (Eerola et al., 2013). Scale/emotion associations are found not just in Western music but in the

¹ Institute of Musicology, University of Cologne, Cologne, Germany

² Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, ON, Canada

Corresponding author:

Rie Asano, Department of Systematic Musicology, Institute of Musicology, University of Cologne, Albertus-Magnus-Platz, 50923 Cologne, Germany.
Email: rie.asano@uni-koeln.de



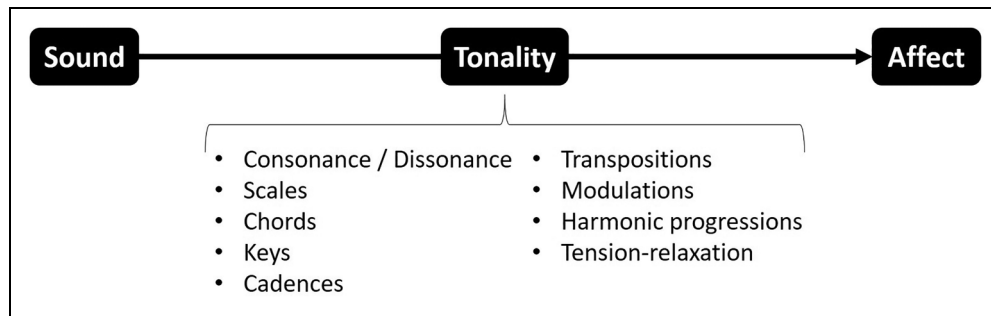


Figure 1. Features of musical tonality.

traditional musics of India, China, Japan, the Middle East, Persia, and Indonesia, among others (Malm, 1996). Tension-relaxation patterns and scale/emotion associations make music into a powerful device for communicating emotion, for example in the accentuation of emotional meanings in songs with words and in narrative works like ballets and films. Tonality is thus a means of *conveying* emotional meanings in music. It is but one mechanism among many by which the arts are able to create cognitive representations of emotion and serve as expressive objects (Davies, 1994, 2001; Hatten, 2018; Kivy, 1980, 1990).

This notion of expression is supported by the model of musical communication developed by Juslin and Sloboda (2013), according to which musical communication proceeds by means of a sender (i.e., a composer or performer) conveying emotional meanings through musical features to a receiver (i.e., a listener), who decodes the conveyed emotional meanings. The receiver perceives the conveyed emotions, but does not necessarily feel them. The relationship between musical features and perceived emotions can be either iconic (i.e., similarity-based), indexical (i.e., association-based), or symbolic (i.e., convention-based). Tonality is one of the musical features that conveys emotional meanings in musical communication, mainly based on the internal, syntactic relationship between pitches, as determined by conventional rules.

While musical intervals can be found in speech (Chow & Brown, 2018; Patel, 2008; Steele, 1775), they are not semantically salient in either the production or perception of speech, even in the case of so-called tone languages. By contrast, discrete intervals – whether in a melodic or harmonic context – are the central feature of music as a cognitive phenomenon. They are the foundation of music’s domain specificity (Brown, 2022; Lerdahl, 2013; Podlipniak, 2013, 2017). Given this specificity, it would stand to reason that there should be some degree of neural specificity for tonality in the human brain, since there is nothing analogous to a musical scale or harmonic progression in speech, let alone in non-acoustic domains. A basic conundrum for the neuroscience of music is that, despite there being ample evidence for music’s domain specificity at the cognitive level, there is minimal evidence for the neural specificity of music when compared with other acoustic functions like speech (Peretz et al., 2015).

Much work on the neuroscience of music has focused on the auditory association cortex of the posterior superior temporal gyrus (pSTG, planum temporale), which has a well-established role in the processing of pitched sounds in music, speech, and environmental sounds (Zatorre et al., 1994; Zatorre et al., 1992). Activations in the right pSTG and inferior frontal gyrus (IFG) are reported in the processing of tonal-harmonic structures (Bianco et al., 2016; Koelsch et al., 2002; Koelsch et al., 2005), and the right arcuate fasciculus connecting these structures shows aberrant structural properties in people with music-processing deficits (Loui et al., 2009). Much controversy has reigned over whether there is any specificity for music in the pSTG. Some models argue for a lateralization of function here, whereby the right pSTG has a preference for musical sounds, compared to the left pSTG for speech sounds (Zatorre et al., 2002). More-recent studies have described an area in the mid-STG bilaterally that is proposed to be specific for music (although this specificity is described by the authors as “weak”), as demonstrated by its responsiveness to a diverse set of musical samples, but lower sensitivity to similar non-musical acoustic stimuli, such as speech and environmental sounds (Norman-Haignere et al., 2015). This area shows moderate responsiveness to drum music that lacks any pitch percept (Boebinger et al., 2021), which raises questions about whether such an area would be a reasonable candidate for being a tonal center in the human brain.

Another brain area that has shown some evidence of music specificity is the anterior part of the STG (aSTG) at the temporal pole (Brodmann areas 22 and 38). The aSTG is considered as part of the ventral auditory stream that responds to phrase-level structures in both music and speech, compared to elemental units in both domains, which are processed more posteriorly in the pSTG. In the case of speech, the aSTG responds more to sentences than to individual words or phonemes (DeWitt & Rauschecker, 2012). For music, Brown et al.’s (2004) production study of singing found that the aSTG was the main area present when melodic singing was contrasted with monotone chanting. Angulo-Perkins et al. (2014) and Angulo-Perkins and Concha (2019) directly compared the perception of song to the perception of speech, and found evidence of music specificity in the aSTG bilaterally. This

effect was stronger in trained musicians than in non-musicians (see also Boebinger et al., 2021).

Finally, another candidate for a music-specific area is the IFG pars orbitalis in Brodmann area (BA) 47 at the neuroanatomical interface between the ventral IFG, anterior insula, and orbitofrontal cortex (OFC). This area is activated in studies that contrast intact music with scrambled versions of the same music (Fedorenko et al., 2012; Levitin & Menon, 2003). It is also present in studies that have looked at tonality-specific processes such as harmonic progressions (Koelsch et al., 2005), cadences (Seger et al., 2013), and tonal tension (Lehne et al., 2014). Moreover, BA 47, together with the posterior superior temporal sulcus, is associated with music-processing deficits (Mandell et al., 2007). Functionally, this area seems to interface emotion with semantic processing (Belyk et al., 2017). This is relevant for music since tonality is used as a means of conveying the emotional contents of the musical object. Therefore, it functions analogously to prosody and emotional language in speech. In fact, BA 47 is a key area for the processing of affective speech prosody. Belyk and Brown's (2014) activation likelihood estimation (ALE) meta-analysis of 19 studies of affective prosody found bilateral activations in BA 47 and the anterior insula just dorsal to the coordinates reported in both of the studies comparing intact music to scrambled versions of it (Fedorenko et al., 2012; Levitin & Menon, 2003). Importantly, BA 47 was the main brain region in the meta-analysis that distinguished affective prosody from linguistic prosody, suggesting that this area processes, at least in part, the acoustic perception of emotion.

In order to examine the neural basis of tonal processing in music, as well as to explore potential neural specificity for music in the human brain, we carried out a voxel-based meta-analysis of 20 published functional magnetic resonance imaging (fMRI) studies of tonal processing in music using activation likelihood estimation (ALE) meta-analysis. The set of studies that was examined in this analysis had a strong leaning towards harmonic progressions in chordal samples, rather than scale structure in monophonic melodies. We predicted that significant areas of concordance across this corpus of studies would include the pSTG, aSTG, and IFG pars orbitalis, with an emphasis on the right hemisphere.

Methods

Search query and inclusion criteria

We searched the PubMed and Google Scholar databases for published fMRI and positron emission tomography (PET) studies using the search terms “fMRI”, “PET”, and “music” along with the following terms: tonal/tonality, chord, grammatical/grammar, syntactic/syntax, melody, harmonic/harmony, tension, musical structure, tonal structure, melodic structure, and harmonic structure. The reference sections of the retrieved publications were searched for additional studies. Figure 2 shows the article screening procedure. The database search was conducted on May 19, 2021 using Publish or Perish (<https://harzing.com/resources/publish-or-perish>).

The inclusion criteria for studies that were contained in the meta-analysis were as follows: 1) that functional brain scanning was performed using either fMRI or PET, thereby excluding studies using electroencephalography, magnetoencephalography, functional near infrared spectroscopy, structural imaging techniques, and resting-state functional connectivity; 2) that the papers reported activation foci in the form of standardized stereotaxic coordinates in either Talairach space or Montreal Neurological Institute (MNI) space (excluding, for example, Minati et al., 2008); 3) that results from the entire scanned brain volume were reported, thereby excluding studies that had partial brain coverage, that reported activation data for only specific areas (e.g., Mueller et al., 2011), or that only reported region-of-interest analyses; 4) that the papers reported the results as standard subtraction analyses, thereby excluding studies using methods like independent components analysis (e.g., Schmithorst, 2005), although we did include a regression analysis of felt tonal tension from Lehne et al. (2014) and one of expectancy violations in cadences from Seger et al. (2013); 5) that the participants were healthy adults, thereby excluding studies using clinical populations and healthy non-adults; and 6) that the study examined key features of tonal processing in music, as shown in Figure 1. The majority of studies in the meta-analysis looked at aspects of harmony processing, rather than melody processing. We excluded studies that performed direct comparisons between the major and minor modes, since their focus was more on the processing of emotional valence than on tonal processing per se (Green et al., 2008; Khalifa et al., 2005; Mizuno & Sugishita, 2007). The included studies covered a combination of passive listening tasks and active discrimination tasks. The participants across the set of studies were a roughly equal combination of musicians and non-musicians, either within- or between-study. We did not examine musical training as a variable in the ALE analysis.

In order to develop a consistent approach to experiment selection, we developed three selection rules for the directionality of the contrasts. 1) For studies that compared tonal with atonal sequences, we selected the “tonal vs. atonal” directionality. Because atonal sequences, in contrast to tonal sequences, lack a tonal center and other related central components of tonality – such as key, scale, and (harmonic) progressions – we argue that tonal processing is less pronounced for atonal sequences than for tonal sequences. The tonal vs. atonal contrast should thus be sensitive to tonal processing. 2) For those studies that compared regular musical sequences with modified, non-typical, incongruent versions of them, we selected the polarity of “irregular vs. regular”, rather than the reverse. By “irregular” musical sequences, we mean sequences that allow listeners to build up a tonal context in the same manner that they do for regular sequences, but that introduce an expectancy violation through an out-of-context tone or chord. Irregular sequences differ from atonal sequences in that they are based on a tonal center and

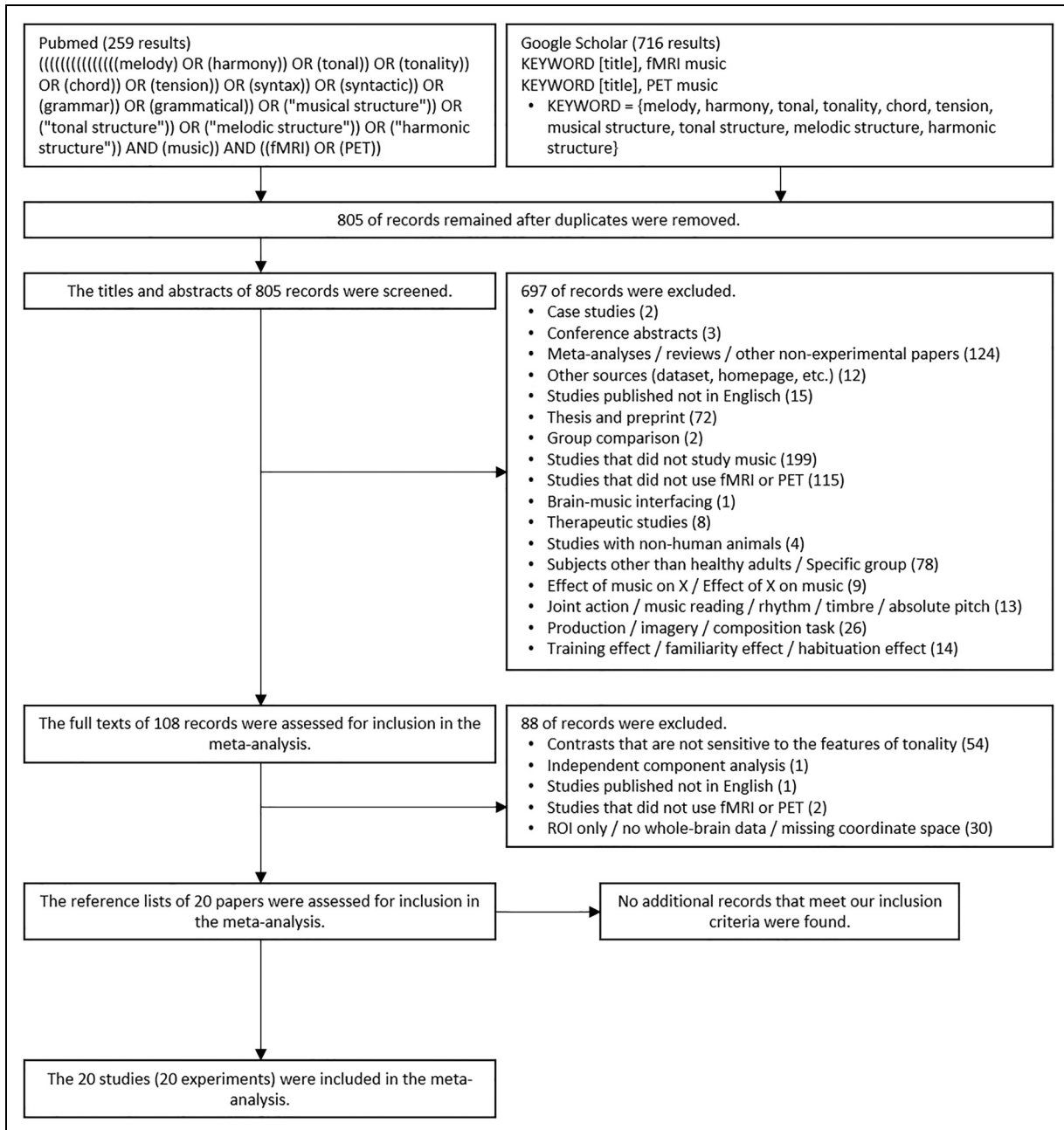


Figure 2. Flowchart of the article screening procedure.

other central components of tonality, with the exception of the out-of-context element that introduces the expectancy violation. In processing irregular sequences, listeners actively integrate the out-of-context element into the established tonal context or establish a new tonal context. Thus, irregular sequences should be stronger elicitors of activation than regular sequences within the same basic music network. Many studies presented the “irregular vs. regular” polarity as the basis for their experimental design using this rationale. 3) For those few studies that compared intact music with scrambled versions of that music, we selected the polarity of “intact music vs. scrambled music”. Because the scrambled music used in these

experiments lacks the central components of tonality, such as key and harmony (Fedorenko et al., 2012), and since it disrupts “navigation through tonal and key spaces” (Levitin & Menon, 2003, p. 2144), we reasoned that tonal-processing regions of the brain would be activated more strongly in the “intact music” condition than in the “scrambled music” condition. No papers reported the reverse contrast alone, and so this rule did not create any complications in experiment selection.

It is important to note that most of the published studies that met our inclusion criteria contained multiple closely related contrasts using a small set of conditions. The meta-analytic practice of selecting multiple experiments from a

Table 1. Listing of the experiments included in the meta-analysis.

Reference	Musical Domain	Main feature	Selected Experiment
Koelsch et al. (2002)	harmony	modulation	modulation > in-key
Levitin and Menon (2003)	music (harmonic)	harmonic progression	normal > scrambled
Tillmann et al. (2003)	harmony	harmonic progression	unrelated consonant > related consonant
Koelsch et al. (2005)	harmony	harmonic progression	irregular chord > regular chord
Tillmann et al. (2006)	harmony	harmonic progression	less-related chord > related chord
Durrant et al. (2007)	harmony	key	key change > no key change
Foster and Zatorre (2010)	melody	transposition	transposed melody > simple melody
Fujisawa and Cook (2011)	harmony	cadence	cadence > white noise
Schulze et al. (2011)	music	scale	tonal sequence > atonal sequence
Fedorenko et al. (2012)	music	harmonic progression	intact music > scrambled music
Foster et al. (2013)	melody	transposition	transposed melody > simple melody
Oechslin et al. (2013)	harmony	harmonic progression	main effect of transgression
Seger et al. (2013)	harmony	cadence	regressor of expectancy violation
Lehne et al. (2014)	music (harmonic)	tonal tension	tension regressor
Spada et al. (2014)	melody/harmony	harmonic progression	altered melody > correct melody
Musso et al. (2015)	harmony	harmonic progression	structural deviant > baseline
Bianco et al. (2016)	harmony	harmonic progression	incongruent > congruent
Cheung et al. (2018)	melody	scale	ungrammatical > grammatical
Tsai and Li (2019)	harmony	transposition	tonality-changed > tonality-unchanged
Li et al. (2021)	harmony	harmonic progression	diatonic sequence > atonal sequence

The references are listed chronologically. All of the included studies are fMRI studies since no PET studies met the inclusion criteria.

given study has the risk of creating duplicate results that artificially increase the concordance of the activated regions from these studies (Müller et al., 2018). In order to avoid this problem, we limited our selection to one experiment per published paper. The means of selecting that experiment was based on the selection rules mentioned above. In general, the experiment that was selected from the two or more closely related experiments in a single study was the one that best matched the tasks in the other studies, without any consideration for the results themselves. The full set of experiments is shown in Table 1. The final meta-analysis included 20 experiments (201 foci, 399 participants) from 20 published fMRI studies. This surpasses the threshold number of experiments required to carry out a valid ALE meta-analysis (Eickhoff et al., 2016; Müller et al., 2018).

ALE meta-analysis

Activation likelihood estimation (ALE) meta-analysis is a coordinate-based statistical method to look for concordant areas of activation across a set of neuroimaging studies (Turkeltaub et al., 2002). Each focus of activation is modeled as a three-dimensional Gaussian probability distribution whose width is determined by the size of the subject group so as to reflect increasing certainty with increasing sample size (Eickhoff et al., 2009). Maps of activation likelihoods are created for each experiment by taking the maximum probability of activation at each voxel. A random-effects analysis tests for the convergence of activations across studies against a null hypothesis of spatially independent brain activations.

All analyses were performed using GingerALE 3.0.2 (www.brainmap.org/ale) according to standard methods

(Eickhoff et al., 2009, 2016; Eickhoff et al., 2012; Müller et al., 2018). Talairach coordinates were converted to MNI coordinates within GingerALE. The meta-analysis was performed as 5,000 threshold permutations using a cluster-level, family-wise error threshold of $p < 0.05$ and a cluster-forming threshold of $p < 0.001$. The ALE scores reported in Table 2 in the Results section are a reflection of the effect sizes reported in standard meta-analyses outside of the neuroimaging field (Eickhoff et al., 2012). The ALE results were registered onto an MNI-normalized template brain using Mango 4.1 (ric.uthscsa.edu/mango).

Results

Figure 3 presents the meta-analysis results for the 20 experiments from 20 published fMRI studies. The MNI coordinates of the ALE clusters are listed in Table 2. The strongest ALE foci occurred as two adjacent clusters in the right frontal lobe at the junction of the IFG pars orbitalis, ventral anterior insula, and OFC in Brodmann areas 47 and 13. One additional cluster occurred in the nearby IFG pars opercularis in BA 44/45 in the right frontal operculum. Additional clusters were seen in the left aSTG (BA 22), the right dorsolateral prefrontal cortex (BA 46/9), and the right superior frontal gyrus (BA 8). No clusters were found in the pSTG at the current threshold. However, one right pSTG cluster appeared when the cluster-forming threshold was reduced to $p < 0.01$ (data not shown). The coronal slice in the lower half of Figure 3 demonstrates the frontal-lobe ALE clusters.

We carried out an additional qualitative analysis of the coordinate tables of the Results section of each paper with regard to general anatomical regions associated with music processing. We did this since a given functional

Table 2. MNI coordinates of the ALE clusters in the meta-analysis.

		BA	x	y	z	ALE	Cluster size (mm ³)
1	Right IFG (pars orbitalis), anterior insula	13	34	24	0	0.026	5,768
			48	16	0	0.020	
		47/13	46	14	-8	0.016	
		47	50	20	-14	0.013	
		44/45	52	18	8	0.023	
	40	20	8	0.019			
	Right DLPFC	46/9	48	20	26	0.016	
2	Left aSTG	22	-52	8	-10	0.024	2,512
			-54	10	0	0.016	
3	Right dorsal IFG	45/46	46	34	2	0.019	1,344
4	Right superior frontal gyrus	8	2	24	54	0.017	888
			6	30	48	0.013	

Stereotaxic coordinates are presented in millimeters along the left-right (x), anterior-posterior (y), and superior-inferior (z) axes. The “ALE” column provides the ALE score for each cluster. Abbreviations: aSTG, anterior part of the superior temporal gyrus; BA, Brodmann area; DLPFC, dorsolateral prefrontal cortex; IFG, inferior frontal gyrus.

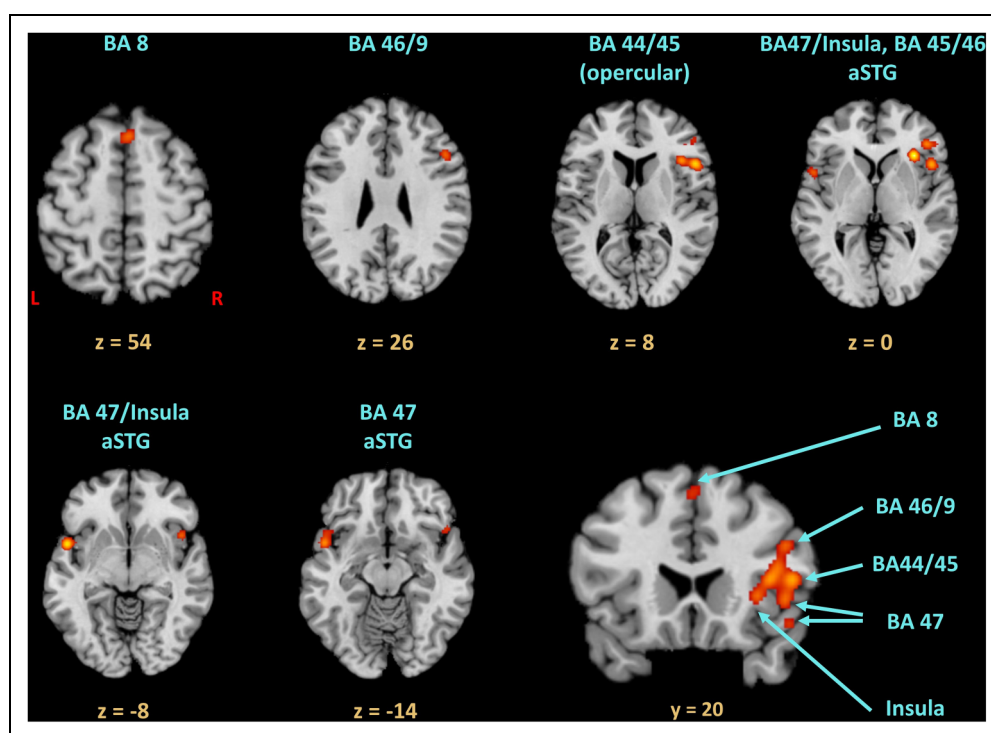


Figure 3. ALE clusters for the tonality meta-analysis. The MNI z level is indicated below each of the slices, except for the coronal slice in the bottom row, which indicates the MNI y level. The left side of the slice (L) is the left side of the brain. The coronal slice shows the five frontal-lobe ALE clusters in the analysis. Abbreviations: aSTG, anterior part of the superior temporal gyrus; BA, Brodmann area.

region may be common across studies but not appear as an ALE cluster in the meta-analysis if the activation locations are not sufficiently overlapping at the voxel level. This analysis revealed that right BA 47, and/or the adjacent ventral anterior insula, was reported in 70% of the publications. This was followed by the right IFG pars opercularis (BA 44/45) at 45%, left aSTG (BA 22/38) at 30%, and the right pSTG (BA 22) at 25%. Right BA 47, and/or the adjacent ventral anterior insula, was not only the strongest ALE cluster at the voxel level, but also the most frequent anatomical region to be reported in the Results sections of these publications.

Discussion

We used ALE meta-analysis to identify brain areas that mediate tonal processing in music across 20 fMRI experiments, with an emphasis on harmonic processing. The results revealed a set of right frontal-lobe areas encompassing BA 47 (and the adjacent insula), the opercular parts of BA 44/45, the DLPFC in BA 46/9, and the superior frontal gyrus in BA 8. The only left-hemisphere cluster was located in the aSTG. An analysis of the primary publications showed that right BA 47 and/or the adjacent ventral insula were reported in 70% of the publications, suggesting

that this area plays a central role in tonal processing for music, most especially harmonic processing, which dominated the studies that were included in the ALE analysis.

How specific are these clusters for music?

The Introduction discussed the fact that music, despite having a significant number of domain-specific cognitive features related to tonality, has shown minimal evidence of neural specificity in neuroimaging studies. Are the clusters of our tonality meta-analysis specific to music and to hierarchical tonal functions like scale structure and harmonic progressions, or are they shared with other functions? As mentioned in the Introduction, Belyk and Brown (2014) carried out an ALE meta-analysis of studies of affective speech prosody, in other words studies in which people had to discriminate the emotions conveyed in spoken utterances based not on lexical content but on prosodic cues related to vocal pitch, loudness, and tempo. They observed bilateral ALE clusters in BA 47. The peaks of the right-hemisphere clusters were located at the posterior ventral portion of BA 47 bordering on the insula, proximate to the peaks in the tonality meta-analysis. These authors also found prosody peaks in right BA 44/45, the right DLPFC, and the left aSTG proximate to the tonality peaks. This suggests that Cluster 1 – including right BA 47/insula, right BA 44/45, and right DLPFC – as well as Cluster 2, including the left aSTG, may not be specific to music but might potentially be shared between music and speech prosody, and might thus be responsive to commonly-used parameters like pitch height and pitch contour, but with no specificity for discrete intervals and musical scales. However, it is important to point out that we have not carried out a statistical comparison between tonal processing in music and affective speech prosody, since this was not within the intended scope of the study. Merrill et al. (2012) carried out a comparative analysis of the prosodic aspect of speech and pitch processing in vocal music in a passive-listening study. They found a lateralization effect such that left BA 47 was associated with speech prosody, and right BA 47 with pitch processing in music.

The largest difference between the present analysis and Belyk and Brown's (2014) prosody results is that the prosody ALE gave a highly bilateral activation profile, whereas the tonality ALE gave a strongly right-lateralized effect. In fact, left BA 47 was the brain region that most distinguished affective speech prosody from linguistic prosody, suggesting a connection with the acoustic correlates of emotion. In Angulo-Perkins et al.'s (2014) direct comparison between music and speech, the music>speech contrast did not show a peak in right BA 47/13 (only in the aSTG), but the speech>music contrast showed a peak in *left* BA 47/13, in keeping with the meta-analysis profile of Belyk and Brown (2014). Moreover, the affective speech prosody meta-analysis did not display any peaks in right BA 45/46 or right BA 8 resembling Clusters 3 and 4 from our tonality meta-analysis. These two regions are suggested to relate to working memory and attention processing for melodic and harmonic

sequences (Brown & Martinez, 2007; Koelsch et al., 2005; Koelsch, Fritz, v. Cramon, Müller, & Friederici, 2006; Oechslin et al., 2013). At present, there is insufficient information to argue that these regions are music-specific. However, this should be explored in future fMRI studies on tonal processing.

To the extent that there may indeed be similarities between tonality and affective speech prosody, how can such similarities be explained? We speculate that one function that could unify tonal processing in music with affective prosody in speech is what we will refer to as “affective semantics”. In contrast to lexical semantics – where words signify certain categories of concepts, such as object-concepts and action-concepts – affective semantics is about conveying *emotional meanings* during communication. For example, musical scales are used connotatively to convey emotional valence during musical communication (Huron, 2008; Parncutt, 2014), and they do so in a categorical manner, unlike the coding of emotional intensity, which occurs along a graded continuum (Eerola et al., 2013). Likewise, harmonic progressions are able to convey a sense of cycling between tension and relaxation (Jackendoff & Lerdahl, 2006; Koelsch, 2011b). While tonal devices such as these are intramusical (Meyer, 1956), they can also be used extramusically to refer phenomena beyond the music itself (Cross & Tolbert, 2016; Juslin & Sloboda, 2013; Koelsch, 2011b), such as in music's use in film narratives (Cohen, 2013; Gorbman, 1987).

It is important to note that a brain system for affective semantics should preferentially process the *conveyed* emotions of the musical object, rather than the emotions that people themselves feel in response to music listening. While our meta-analysis was not based on the analysis of emotion per se (but rather tonal processing), it is interesting to compare our results to those of Koelsch's (2020) ALE meta-analysis of the brain areas activated when people experience felt emotions in response to music listening, for example when music is used “to evoke joy, sadness, fear, tension, frissons, surprise, unpleasantness, or feelings of beauty” (p. 1). The clusters reported by these two analyses are almost completely non-overlapping. BA 47/13 and BA 44 did not appear as ALE clusters in Koelsch's analysis. Instead, a large number of limbic areas not seen in the current analysis were present, most of them associated with the experience of felt emotions, compared to perception of conveyed emotions in communicative media. These included the amygdala, hippocampus, striatum, anterior cingulate cortex, mid-cingulate cortex, OFC, and various parts of the auditory cortex bilaterally. Hence, a comparison between tonal processing in the current meta-analysis and felt emotions in Koelsch's meta-analysis shows almost no overlap. This contrasts with the striking parallel between the tonality ALE and the right-hemisphere peaks in Belyk and Brown's (2014) meta-analysis of affective speech prosody. This observation reinforces our speculation that tonality and speech prosody might share a deep underlying connection with one another via affective semantics and the conveyance of emotions in acoustic communication media. This affective-

semantic function of conveying emotion seems to be neurally distinct from the experience of felt emotions in response to music, as shown by the comparison between the meta-analyses devoted to affective speech prosody (Belyk & Brown, 2014) and music-evoked emotions (Koelsch, 2020).

The current speculation about parallels between tonal processing in music and affective prosody in speech leads to two possible hypotheses regarding putative neural specificity for music in the human brain. The first hypothesis is that there is no neural specificity for music, and that tonality-specific functions like musical scales and harmonic progressions are co-localized with non-tonal functions like speech prosody in regions such as right BA 47/13. The second hypothesis is that music's neural specificity lies downstream of BA 47/13 in the brain, but that BA 47/13 itself encodes prosodic features that are shared between music and speech and that do not distinguish between them, features such as melodic contour, pitch register, loudness, and/or tempo. Such a view would be consistent with evolutionary proposals dating back to the Enlightenment that music and speech co-evolved from a joint prosodic precursor involved in vocal communication (Brown, 2000, 2017; Mithen, 2005; Rousseau, 1781; Wallaschek, 1891). If this proposal is correct, then BA 47/13 might be a neural remnant of this evolutionary process, in which case a tonality-specific region downstream of BA 47/13 might be discoverable in future fMRI studies of tonal processing in music, especially ones that employ speech prosody as a comparison condition.

While a number of fMRI studies have directly contrasted the major and minor scales with one another (Green et al., 2008; Khalifa et al., 2005; Mizuno & Sugishita, 2007), they have not contrasted scales with non-scales, which is the type of design that is necessary to identify tonality-specific areas in the brain. In a scale, pitches are organized in relation to a tonal center (Lerdahl, 2001), while this is not the case for a non-scale pitch sequence. Contrasting a scale with a non-scale condition is thus one way to reveal tonality-specific brain regions. Studies that have directly contrasted music with non-musical functions like speech have revealed the importance of the aSTG bilaterally for music (Angulo-Perkins et al., 2014; Angulo-Perkins & Concha, 2019).

While the current work was in preparation, a meta-analysis of 50 neuroimaging experiments of music listening was published by Chan and Han (2022). The focus of this analysis was not on tonality per se or on any aspect of active musical processing, but on passive music listening alone. The strongest cluster in this analysis occurred in the right frontal operculum in the vicinity of BA 44, BA 47, and the anterior insula. While the coordinates of this cluster and its multiple subclusters were different than those reported in the present meta-analysis, they do indicate that tonality areas are activated by passive listening, rather than requiring active discrimination. In contrast to the current results, the analysis of Chan and Han showed that the music-listening network is highly bilateral and that it includes a series of limbic and subcortical brain areas, including the hippocampus, amygdala, cerebellum, basal ganglia, and thalamus. A number of these areas are those

reported in studies of felt emotions in response to music listening (see above). Some of the differences between two meta-analyses stem from methodological differences. Chan and Han, for example, included contrasts against “rest”, which helps explain why their meta-analysis contained more brain areas overall. Their meta-analysis also included studies conducted using more-realistic musical excerpts and of rhythm processing, which were not included in our analysis because of its exclusive focus on tonal processing.

The ventral auditory pathway for affective semantics?

Proposals have been made that the auditory pathways of the brain are organized according to parallel dorsal and ventral streams connecting the frontal and temporal lobes, analogous to a similar two-stream segregation in the visual system (Rauschecker & Scott, 2009; Rauschecker & Tian, 2000). One component of the dorsal auditory pathway in humans is the arcuate fasciculus connecting the pSTG to the posterior IFG (i.e., BA 44/45), which is associated with sequence processing, auditory-motor mapping, and syntax (Bornkessel-Schlesewsky & Schlesewsky, 2013; Friederici, 2012, 2019; Hickok & Poeppel, 2007; Zatorre et al., 2007). The posterior IFG was represented in the current meta-analysis by the ALE cluster in the opercular part of BA 44/45 in the right hemisphere (see Figure 3).

The ventral pathway is more of a categorical system for auditory object recognition that is involved in mapping meaningful information onto communication sounds, both lexical and affective meanings (Hickok & Poeppel, 2007; Schirmer & Kotz, 2006). The ventral pathway consists of projections to the anterior IFG (i.e., BA 45 and 47) from the aSTG via the uncinate fasciculus (UF) and from the pSTG (and the posterior middle temporal gyrus) via the extreme capsule fasciculus (EmC) (Makris & Pandya, 2009; Weiller et al., 2021). The anterior IFG was represented in the current meta-analysis by the ALE cluster in right BA 47, and the aSTG was represented by the cluster in left BA 22 (see Figure 3).

We have argued above that affective semantics might provide a reasonable basis for uniting tonality in music with the prosody of speech, not least since tonality tends to operate using relatively discrete categories, like scale types, keys, chord types, and cadence types. In light of the dual-stream model of auditory processing, we propose that the ventral auditory pathway is a candidate for implementing affective semantics in the brain. Along these lines, Frühholz et al.'s (2015) probabilistic fiber tracking study showed that the processing of affective prosody engages not only the dorsal pathway, but the ventral pathway as well. The left aSTG, as well as the right IFG, are involved in the ventral pathway associated with affective prosody processing. Moreover, Belyk et al.'s (2017) kernel density meta-analysis of the IFG pars orbitalis (BA 47) found a lateralization effect whereby the left orbitalis is associated with lexical meanings, whereas the right orbitalis is associated with affective meanings, including affective speech prosody. Hartwigsen et al.'s (2019) coactivation-based parcellation

study suggested a social and emotional role for the right anterior IFG, including BA 47. In addition, Goodkind et al.'s (2012) voxel-based morphometry study found that BA 47 is central to dynamically tracking emotional valence.

From a musical standpoint, the right ventral pathway is associated with acquired amusia (Sihvonen et al., 2017), and the right IFG orbitalis shows anomalies in subjects with congenital amusia (Hyde et al., 2006; Hyde et al., 2007; Hyde et al., 2011). In addition, amusic individuals show impaired performance in explicitly judging the emotional prosody of short vocal samples based on pitch and spectro-temporal parameters (Pralus et al., 2019). Aprosodia and amusia are jointly associated with an abnormality in the right ventral pathway (Sihvonen et al., 2022). Overall, the ventral auditory pathway might not merely be a semantic pathway, but a system that also encodes affective semantics, especially in the right hemisphere. Given the strong parallel between speech prosody and tonal processing in the ventral auditory pathway, we predict that there should be areas downstream of BA 47 that show specificity for tonal processing in music but that have not yet been characterized as such (although see Janata et al., 2002).

The right ventral pathway seems to have a hierarchical organization, in which pitch information processed in the pSTG is projected to the aSTG, integrating emotionally significant cues into a unit, and then to areas processing affective semantics in right BA 47, the ventral insula, and the adjacent part of the OFC, as conveyed through the EmC and/or UF (Schirmer & Kotz, 2006). Studies of functional connectivity support a connection between the anterior temporal lobe and both BA 47 and the OFC (Jung et al., 2017). Such a pathway might have right-hemisphere dominance. The volume of the UF shows right-hemisphere lateralization for social/emotional processing, compared to the left for semantic processing in language (Papinutto et al., 2016).

It is important to note that our tonality meta-analysis did not show a cluster in the right temporal lobe, except for a right pSTG cluster that emerged when the cluster-forming threshold was reduced beyond the standard threshold of $p < 0.001$. Nevertheless, studies that have directly contrasted music with speech highlight the importance of the bilateral aSTG for music (Angulo-Perkins et al., 2014; Angulo-Perkins & Concha, 2019), and a study on pitch-based hierarchical structure building reported activation in the right pSTG (Martins et al., 2020). Thus, the specificity of the right temporal lobe for higher-level tonal processing needs further clarification, despite the well-established role of this area in low-level pitch processing for music (Zatorre et al., 1994, 1992).

Implications for comparative research on language and music

We have thus far focused on the relationship between affective prosody and tonality from the perspective of affective semantics. However, the results of the current meta-analysis have additional implications for comparative research on language and music. First, affective semantics is a function

that could be associated with another central component of music and prosody, namely rhythm. Musical rhythm encodes tension-relaxation patterns through tempo, syncopation, and polyrhythm (Pressing, 2002; Trost et al., 2017; Vuust & Witek, 2014). Activation in BA 47 has been reported for the processing of syncopated rhythms (Mayville et al., 2002) and polyrhythms (Vuust et al., 2006; Vuust et al., 2011). Second, beyond affective functions, music and prosody share processes in segmentation, prominence, and coordination (Palmer & Hutchins, 2006). Because speech segmentation and prominence through the modulation of pitch height and/or loudness are associated with the right frontal operculum (BA 44) (Belyk & Brown, 2014), the right IFG cluster in our meta-analysis could relate more strongly to these functions than to affective functions. The role of rhythm should be considered in future research, given the tight relationship between musical rhythm and prosody (Hausen et al., 2013).

The relationship between tonal processing in music and syntactic processing in language has been repeatedly discussed because of their abstract, rule-based, and hierarchical properties (e.g., Asano & Boeckx, 2015; Koelsch, 2011a, 2012; Patel, 2003, 2008, 2013). Because hierarchical processing in language, music, and action all engage Broca's region, including BA 44 and 45, this region has been suggested to be a domain-general hierarchical processor (Fitch & Martins, 2014). In a similar vein, cognitive control has been proposed as a shared mechanism for hierarchical processing in language, music, and action (Jeon, 2014; Slevc & Okada, 2015). Hierarchical predictive processing, i.e., processing expectancy and expectancy violations, is also an important candidate mechanism shared in language and music (Koelsch et al., 2019; Rohrmeier & Koelsch, 2012). From these perspectives, the role of the frontal operculum in both tonal processing and prosody might be interpreted in terms of domain-general hierarchical processing (see also Heffner & Slevc, 2015 for discussions about hierarchical structure of music and prosody). Thus, future research on the relationship between music and prosody, especially a direct quantitative comparison, may contribute to clarifying the relationship between language and music in terms of hierarchical processing (Chen et al., 2021) and thus inform the current domain-generality vs. -specificity discussion in cognitive neuroscience in an important way (Asano et al., 2022).

Limitations

There are a number of significant limitations in the present study. 1) A relatively small number of published studies was available for the analysis, although this number exceeded the 17-experiment threshold required for running a statistically valid ALE meta-analysis (Eickhoff et al., 2016; Müller et al., 2018). 2) The included studies were very heterogeneous in musical focus and experimental design. The tasks were very diverse, covering passive music listening and active discrimination, and doing so for

harmonic progressions, cadences, transposition, and the like. 3) Related to the last point, the polarity of the contrasts was variable across papers in the literature. In the end, we selected the contrasts that would maximally represent tonal processing, as mentioned in the Methods section. What has yet to be done is a study that directly compares a tonal condition against a non-tonal condition for music in order to identify brain areas specific for tonal processing. Several studies have compared the major and minor scales directly, but such studies do not permit an assessment of music specificity since both conditions involve the same tonal process. 4) In addition, there is a need for experimental approaches to tonal processing that look beyond expectancy violations per se, since studies using this type of design comprised fully 40% of the studies in the meta-analysis. 5) Very few studies have looked at monophonic melodies and the processing of musical scales independent of chords and/or harmonic sequences. Because the literature on tonal processing is skewed towards harmonic contexts, it is not known to what extent tonal-processing areas are activated in monophonic contexts that reflect scale effects in melodies. That should be a priority of future work.

6) Finally, a limitation of this work that has nothing to do with the included experiments or the analytical approach is the complexity of the neuroanatomy in the frontal-lobe region in which we identified the most significant ALE clusters. This is in a diverse region that interfaces the IFG pars orbitalis (BA 47), the ventral part of the anterior insula (BA 13), and the posterior part of the OFC (what some sources call BA 12), all associated with social and emotional functions (Kurth et al., 2010; Rolls et al., 2020; Wojtasik et al., 2020). Adding to the anatomical complexity of this frontal-lobe region is the fact that the anterior tip of the temporal lobe is directly posterior to it, most notably the aSTG. Hence, we could imagine that an activation peak could be assigned to different lobes depending on the template brain that was used for registration and the dimensions of the temporal pole vis-à-vis the IFG pars orbitalis, insula, and OFC. Future research on affective semantics should take this complexity into account to enable a more fine-grained analysis of the relationship between music, prosody, and other social and emotional functions.

Conclusions

We carried out an ALE meta-analysis of 20 functional MRI studies of tonal processing in music, with an emphasis on harmonic processing, and identified a set of right frontal-lobe areas, including the IFG pars orbitalis (BA 47/13), the frontal operculum (BA 44/45), and the DLPFC (BA 46/9). These areas were found to be very similar to previously reported meta-analytic peaks for affective speech prosody, thereby suggesting that these areas, despite mediating a complex level of musical processing, may not be specific to music, but may instead mediate affective semantics in the conveyance of communication sounds, whether these sounds be tonal/intervallic like music or non-tonal

like speech. Future fMRI studies will need to explore potential music specificity in the brain beyond the areas described here, not least through direct comparisons between music processing and affective speech prosody.

Acknowledgements

Special thanks to the editor, two reviewers, and Uwe Seifert for helpful comments on an earlier version of this manuscript.

Action Editor

Daniela Sammler, Max Planck Institute for Human Cognitive and Brain Sciences.

Peer review

Vincent Cheung, Institute of Information Science Academia Sinica.

Renzo Torrecuso, Max Planck Institute for Human Cognitive and Brain Sciences, NMR.

Author Contributions

RA conceived of the study and carried out the initial analysis. SB and VL contributed to the subsequent study design and data analysis. RA and SB wrote the manuscript. All authors reviewed and edited the manuscript and approved the final version of the manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by a grant to RA from MEXT/JSPS Grant-in-Aid for Scientific Research on Innovative Areas #4903 (Evolinguistics) [grant number JP17H06379] and to SB from the Natural Sciences and Engineering Research Council (NSERC) of Canada (grant number RGPIN-2020-05718).

ORCID iD

Rie Asano  <https://orcid.org/0000-0002-6901-7885>

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), e0222796. <https://doi.org/10.1371/journal.pone.0222796>
- Asano, R., & Boeckx, C. (2015). Syntax in language and music: What is the right level of comparison? *Frontiers in Psychology*, *6*, 942. <https://doi.org/10.3389/fpsyg.2015.00942>

- Asano, R., Boeckx, C., & Fujita, K. (2022). Moving beyond domain-specific versus domain-general options in cognitive neuroscience. *Cortex*, *154*, 259–268. <https://doi.org/10.1016/j.cortex.2022.05.004>
- 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., Brown, S., Lim, J., & Kotz, S. A. (2017). Convergence of semantics and emotional expression within the IFG pars orbitalis. *NeuroImage*, *156*, 240–248. <https://doi.org/10.1016/j.neuroimage.2017.04.020>
- Bianco, R., Novembre, G., Keller, P. E., Kim, S.-G., Scharf, F., Friederici, A. D., & Sammler, D. (2016). Neural networks for harmonic structure in music perception and action. *NeuroImage*, *142*, 454–464. <https://doi.org/10.1016/j.neuroimage.2016.08.025>
- Boebinger, D., Norman-Haignere, S., 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>
- Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2013). Reconciling time, space and function: A new dorsal-ventral stream model of sentence comprehension. *Brain and Language*, *125*(1), 60–76. <https://doi.org/10.1016/j.bandl.2013.01.010>
- 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*, 1894. <https://doi.org/10.3389/fpsyg.2017.01894>
- Brown, S. (2022). *The unification of the arts: A framework for understanding what the arts share and why*. Oxford University Press.
- 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>
- Brown, S., Martinez, M. J., Hodges, D. A., Fox, P. T., & Parsons, L. M. (2004). The song system of the human brain. *Cognitive Brain Research*, *20*(3), 363–375. <https://doi.org/10.1016/j.cogbrainres.2004.03.016>
- Chan, M. M. Y., & Han, Y. M. Y. (2022). The functional brain networks activated by music listening: A neuroimaging meta-analysis and implications for treatment. *Neuropsychology*, *36*, 4–22. <https://doi.org/10.1037/neu0000777>
- Chen, X., Affourtit, J., Ryskin, R., Regev, T. I., Norman-Haignere, S., Jouravlev, O., & Fedorenko, E. (2021). The human language system does not support music processing. *BioRxiv*. <https://doi.org/10.1101/2021.06.01.446439>
- Cheung, V. K. M., Meyer, L., Friederici, A. D., & Koelsch, S. (2018). The right inferior frontal gyrus processes nested non-local dependencies in music. *Scientific Reports*, *8*(1), 3822. <https://doi.org/10.1038/s41598-018-22144-9>
- Chow, I., & Brown, S. (2018). A musical approach to speech melody. *Frontiers in Psychology*, *9*, 247. <https://doi.org/10.3389/fpsyg.2018.00247>
- 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.
- Cross, I., & Tolbert, C. (2016). Music and meaning. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford handbook of music psychology* (2nd ed., pp. 33–46). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198722946.013.7>
- Davies, S. (1994). *Musical meaning and expression*. Cornell University Press.
- 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.
- DeWitt, I., & Rauschecker, J. P. (2012). Phoneme and word recognition in the auditory ventral stream. *Proceedings of the National Academy of Sciences*, *109*(8), E505–E514. <https://doi.org/10.1073/pnas.1113427109>
- Durrant, S., Hardoon, D. R., Miranda, E. R., Shawe-Taylor, J., Brechmann, A., & Scheich, H. (2007). Neural correlates of tonality in music. In *Proceedings of NIPS Workshop on Music, Brain and Cognition*. Whistler, Canada.
- Eerola, T., Friberg, A., & Bresin, R. (2013). Emotional expression in music: Contribution, linearity, and additivity of primary musical cues. *Frontiers in Psychology*, *4*, 487. <https://doi.org/10.3389/fpsyg.2013.00487>
- Eickhoff, S. B., Bzdok, D., Laird, A. R., Kurth, F., & Fox, P. T. (2012). Activation likelihood estimation meta-analysis revisited. *NeuroImage*, *59*(3), 2349–2361. <https://doi.org/10.1016/j.neuroimage.2011.09.017>
- Eickhoff, S. B., Laird, A. R., Grefkes, C., Wang, L. E., Zilles, K., & Fox, P. T. (2009). Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: A random-effects approach based on empirical estimates of spatial uncertainty. *Human Brain Mapping*, *30*(9), 2907–2926. <https://doi.org/10.1002/hbm.20718>
- Eickhoff, S. B., Nichols, T. E., Laird, A. R., Hoffstaedter, F., Amunts, K., Fox, P. T., & Eickhoff, C. R. (2016). Behavior, sensitivity, and power of activation likelihood estimation characterized by massive empirical simulation. *NeuroImage*, *137*, 70–85. <https://doi.org/10.1016/j.neuroimage.2016.04.072>
- Fedorenko, E., McDermott, J. H., Norman-Haignere, S., & Kanwisher, N. (2012). Sensitivity to musical structure in the human brain. *Journal of Neurophysiology*, *108*(12), 3289–3300. <https://doi.org/10.1152/jn.00209.2012>
- Fitch, W. T., & Martins, M. D. (2014). Hierarchical processing in music, language, and action: Lashley revisited. *Annals of the New York Academy of Sciences*, *1316*(1), 87–104. <https://doi.org/10.1111/nyas.12406>
- 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>
- Foster, N. E. V., & Zatorre, R. J. (2010). A role for the intraparietal sulcus in transforming musical pitch information. *Cerebral Cortex*, *20*(6), 1350–1359. <https://doi.org/10.1093/cercor/bhp199>
- Friederici, A. D. (2012). The cortical language circuit: From auditory perception to sentence comprehension. *Trends in Cognitive Sciences*, *16*(5), 262–268. <https://doi.org/10.1016/j.tics.2012.04.001>

- Friederici, A. D. (2019). Hierarchy processing in human neurobiology: How specific is it? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1789), 20180391. <https://doi.org/10.1098/rstb.2018.0391>
- Frühholz, S., Gschwind, M., & Grandjean, D. (2015). Bilateral dorsal and ventral fiber pathways for the processing of affective prosody identified by probabilistic fiber tracking. *NeuroImage*, 109, 27–34. <https://doi.org/10.1016/j.neuroimage.2015.01.016>
- Fujisawa, T. X., & Cook, N. D. (2011). The perception of harmonic triads: An fMRI study. *Brain Imaging and Behavior*, 5(2), 109–125. <https://doi.org/10.1007/s11682-011-9116-5>
- Goodkind, M. S., Sollberger, M., Gyurak, A., Rosen, H. J., Rankin, K. P., Miller, B., & Levenson, R. (2012). Tracking emotional valence: The role of the orbitofrontal cortex. *Human Brain Mapping*, 33(4), 753–762. <https://doi.org/10.1002/hbm.21251>
- Gorbman, C. (1987). *Unheard melodies: Narrative film music*. Indiana University Press.
- Green, A. C., Bærentsen, 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>
- Hartwigsen, G., Neef, N. E., Camilleri, J. A., Margulies, D. S., & Eickhoff, S. B. (2019). Functional segregation of the right inferior frontal gyrus: Evidence from coactivation-based parcellation. *Cerebral Cortex*, 29(4), 1532–1546. <https://doi.org/10.1093/cercor/bhy049>
- Hatten, R. S. (2018). *A theory of virtual agency for western art music*. Indiana University Press.
- Hausen, M., Torppa, R., Salmela, V. R., Vainio, M., & Särkämö, T. (2013). Music and speech prosody: A common rhythm. *Frontiers in Psychology*, 4, 566. <https://doi.org/10.3389/fpsyg.2013.00566>
- Heffner, C. C., & Slevc, L. R. (2015). Prosodic structure as a parallel to musical structure. *Frontiers in Psychology*, 6, 1962. <https://doi.org/10.3389/fpsyg.2015.01962>
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8, 393–402. <https://doi.org/10.1038/nrn2113>
- 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>
- Hyde, K. L., Lerch, J. P., Zatorre, R. J., Griffiths, T. D., Evans, A. C., & Peretz, I. (2007). Cortical thickness in congenital amusia: When less is better than more. *Journal of Neuroscience*, 27(47), 13028–13032. <https://doi.org/10.1523/JNEUROSCI.3039-07.2007>
- Hyde, K. L., Zatorre, R. J., Griffiths, T. D., Lerch, J. P., & Peretz, I. (2006). Morphometry of the amusic brain: A two-site study. *Brain*, 129, 2562–2570. <https://doi.org/10.1093/brain/awl204>
- Hyde, K. L., Zatorre, R. J., & Peretz, I. (2011). Functional MRI evidence of an abnormal neural network for pitch processing in congenital amusia. *Cerebral Cortex*, 21(2), 292–299. <https://doi.org/10.1093/cercor/bhq094>
- Jackendoff, R., & Lerdahl, F. (2006). The capacity for music: What is it, and what's special about it? *Cognition*, 100, 33–72. <https://doi.org/10.1016/j.cognition.2005.11.005>
- 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 (New York, N.Y.)*, 298(5601), 2167–2170. <https://doi.org/10.1126/science.1076262>
- Jeon, H.-A. (2014). Hierarchical processing in the prefrontal cortex in a variety of cognitive domains. *Frontiers in Systems Neuroscience*, 8, 223. <https://doi.org/10.3389/fnsys.2014.00223>
- Jung, J. Y., Cloutman, L. L., Binney, R. J., & Lambon Ralph, M. A. (2017). The structural connectivity of higher order association cortices reflects human functional brain networks. *Cortex*, 97, 221–239. <https://doi.org/10.1016/j.cortex.2016.08.011>
- Juslin, P. N., & Sloboda, J. A. (2013). Music and emotion. In D. Deutsch (Ed.), *The psychology of music* (3rd ed., pp. 583–645). Academic Press. <https://doi.org/10.1016/B978-0-12-381460-9.00015-8>
- 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.
- Koelsch, S. (2011a). Toward a neural basis of music perception: A review and updated model. *Frontiers in Psychology*, 2, 110. <https://doi.org/10.3389/fpsyg.2011.00110>
- Koelsch, S. (2011b). Towards a neural basis of processing musical semantics. *Physics of Life Reviews*, 8(2), 89–105. <https://doi.org/10.1016/j.plrev.2011.04.004>
- Koelsch, S. (2012). *Brain and music*. Wiley-Blackwell.
- Koelsch, S. (2020). A coordinate-based meta-analysis of music-evoked emotions. *NeuroImage*, 223, 117350. <https://doi.org/10.1016/j.neuroimage.2020.117350>
- Koelsch, S., Fritz, T., Schulze, K., Alsop, D., & Schlaug, G. (2005). Adults and children processing music: An fMRI study. *NeuroImage*, 25(4), 1068–1076. <https://doi.org/10.1016/j.neuroimage.2004.12.050>
- Koelsch, S., Fritz, T., v. Cramon, D. Y., Müller, K., & Friederici, A. D. (2006). Investigating emotion with music: An fMRI study. *Human Brain Mapping*, 27(3), 239–250. <https://doi.org/10.1002/hbm.20180>
- Koelsch, S., Gunter, T. C., v. Cramon, D. Y., Zysset, S., Lohmann, G., & Friederici, A. D. (2002). Bach Speaks: A cortical “language-network” serves the processing of music. *NeuroImage*, 17(2), 956–966. <https://doi.org/10.1006/nimg.2002.1154>
- Koelsch, S., Vuust, P., & Friston, K. J. (2019). Predictive processes and the peculiar case of music. *Trends in Cognitive Sciences*, 23(1), 63–77. <https://doi.org/10.1016/j.tics.2018.10.006>
- Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. Oxford University Press.
- Krumhansl, C. L. (2002). Music: A link between cognition and emotion. *Current Directions in Psychological Science*, 11(2), 45–50. <https://doi.org/10.1111/1467-8721.00165>
- Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial

- representation of musical keys. *Psychological Review*, 89(4), 334–368. <https://doi.org/10.1037/0033-295X.89.4.334>
- Kurth, F., Zilles, K., Fox, P. T., Laird, A. R., & Eickhoff, S. B. (2010). A link between the systems: Functional differentiation and integration within the human insula revealed by meta-analysis. *Brain Structure and Function*, 214(5–6), 519–534. <https://doi.org/10.1007/s00429-010-0255-z>
- Lehne, M., Rohrmeier, M., & Koelsch, S. (2014). Tension-related activity in the orbitofrontal cortex and amygdala: An fMRI study with music. *Social Cognitive and Affective Neuroscience*, 9(10), 1515–1523. <https://doi.org/10.1093/scan/nst141>
- Lerdahl, F. (2001). *Tonal pitch space*. Oxford University Press.
- Lerdahl, F. (2013). Musical syntax and its relation to linguistic syntax. In M. A. Arbib (Ed.), *Language, music, and the brain* (pp. 257–272). The MIT Press.
- Levitin, D. J., & Menon, V. (2003). Musical structure is processed in “language” areas of the brain: A possible role for Brodmann Area 47 in temporal coherence. *NeuroImage*, 20(4), 2142–2152. <https://doi.org/10.1016/j.neuroimage.2003.08.016>
- Li, C.-W., Guo, F. Y., & Tsai, C.-G. (2021). Predictive processing, cognitive control, and tonality stability of music: An fMRI study of chromatic harmony. *Brain and Cognition*, 151(1), 105751. <https://doi.org/10.1016/j.bandc.2021.105751>
- Loui, P., Alsop, D., & Schlaug, G. (2009). Tone deafness: A new disconnection syndrome? *The Journal of Neuroscience*, 29(33), 10215–10220. <https://doi.org/10.1523/JNEUROSCI.1701-09.2009>
- Makris, N., & Pandya, D. N. (2009). The extreme capsule in humans and rethinking of the language circuitry. *Brain Structure and Function*, 213(3), 343–358. <https://doi.org/10.1007/s00429-008-0199-8>
- Malm, W. P. (1996). *Music cultures of the pacific, the near east and Asia* (3rd ed). Prentice Hall.
- Mandell, J., Schulze, K., & Schlaug, G. (2007). Congenital amusia: An auditory-motor feedback disorder? *Restorative Neurology and Neuroscience*, 25(3–4), 323–334.
- Martins, M. J. D., Fischmeister, F. P. S., Gingras, B., Bianco, R., Puig-Waldmueller, E., Villringer, A., & Beisteiner, R. (2020). Recursive music elucidates neural mechanisms supporting the generation and detection of melodic hierarchies. *Brain Structure and Function*, 225(7), 1997–2015. <https://doi.org/10.1007/s00429-020-02105-7>
- Mayville, J. M., Jantzen, K. J., Fuchs, A., Steinberg, F. L., & Kelso, J. A. S. (2002). Cortical and subcortical networks underlying syncopated and synchronized coordination revealed using fMRI. *Human Brain Mapping*, 17(4), 214–229. <https://doi.org/10.1002/hbm.10065>
- 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, 76. <https://doi.org/10.3389/fpsyg.2012.00076>
- Meyer, L. B. (1956). *Emotion and meaning in music*. University of Chicago Press.
- Minati, L., Rosazza, C., D’Incerti, L., Pietrocini, E., Valentini, L., Scaioni, V., & Bruzzone, M. G. (2008). fMRI/ERP of musical syntax: Comparison of melodies and unstructured note sequences. *NeuroReport*, 19(14), 1381–1385. <https://doi.org/10.1097/WNR.0b013e32830c694b>
- Mithen, S. (2005). *The singing neanderthals: The origins of music, language, mind and body*. Weidenfeld & Nicolson.
- Mizuno, T., & Sugishita, M. (2007). Neural correlates underlying perception of tonality-related emotional contents. *NeuroReport*, 18(16), 1651–1655. <https://doi.org/10.1097/WNR.0b013e3282f0b787>
- Mueller, K., Mildner, T., Fritz, T., Lepsien, J., Schwarzbauer, C., Schroeter, M. L., & Möller, H. E. (2011). Investigating brain response to music: A comparison of different fMRI acquisition schemes. *NeuroImage*, 54(1), 337–343. <https://doi.org/10.1016/j.neuroimage.2010.08.029>
- Müller, V. I., Cieslik, E. C., Laird, A. R., Fox, P. T., Radua, J., Mataix-Cols, D., & Eickhoff, S. B. (2018). Ten simple rules for neuroimaging meta-analysis. *Neuroscience and Biobehavioral Reviews*, 84, 151–161. <https://doi.org/10.1016/j.neubiorev.2017.11.012>
- Musso, M., Weiller, C., Horn, A., Glauche, V., Umarova, R., Hennig, J., & Rijntjes, M. (2015). A single dual-stream framework for syntactic computations in music and language. *NeuroImage*, 117, 267–283. <https://doi.org/10.1016/j.neuroimage.2015.05.020>
- 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>
- Oechslin, M. S., Van De Ville, D., Lazeyras, F., Hauert, C. A., & James, C. E. (2013). Degree of musical expertise modulates higher order brain functioning. *Cerebral Cortex*, 23(9), 2213–2224. <https://doi.org/10.1093/cercor/bhs206>
- Palmer, C., & Hutchins, S. (2006). What is musical prosody? In B. H. Ross (Ed.), *Psychology of learning and motivation: Advances in research and theory* (Vol. 46, pp. 245–278). Academic Press.
- Papinutto, N., Galantucci, S., Mandelli, M. L., Gesierich, B., Jovicich, J., Caverzasi, E., & Gorno-Tempini, M. L. (2016). Structural connectivity of the human anterior temporal lobe: A diffusion magnetic resonance imaging study. *Human Brain Mapping*, 37(6), 2210–2222. <https://doi.org/10.1002/hbm.23167>
- Parncutt, R. (2014). The emotional connotations of major versus minor tonality: One or more origins? *Musicae Scientiae*, 18, 324–353. <https://doi.org/10.1177/1029864914542842>
- Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, 6(7), 674–681. <https://doi.org/10.1038/nn1082>
- Patel, A. D. (2008). *Music, language, and the brain*. Oxford University Press.
- Patel, A. D. (2013). Sharing and nonsharing of brain resources for language and music. In M. A. Arbib (Ed.), *Language, music, and the brain* (pp. 329–355). MIT Press.
- Peretz, I., Vuvan, D., Lagrois, M.-É., & Armony, J. L. (2015). Neural overlap in processing music and speech. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 370(1664), 20140090. <https://doi.org/10.1098/rstb.2014.0090>

- Podlipniak, P. (2013). Tonality as one of the “music specific” adaptations. In M. Reybrouck, C. Maeder, A. Helbo, & E. Tarasti (Eds.), *E-proceedings of the XIIIth international congress of musical signification* (pp. 310–315). Louvain-la-Neuve.
- Podlipniak, P. (2017). Tonal qualia and the evolution of music. *Avant*, 8(1), 33–44. <https://doi.org/10.26913/80102017.0101.0002>
- Pralus, A., Fornoni, L., Bouet, R., Gomot, M., Bhatara, A., Tillmann, B., & Caclin, A. (2019). Emotional prosody in congenital amusia: Impaired and spared processes. *Neuropsychologia*, 134, 107234. <https://doi.org/10.1016/j.neuropsychologia.2019.107234>
- Pressing, J. (2002). Black Atlantic rhythm: Its computational and transcultural foundations. *Music Perception*, 19(3), 285–310. <https://doi.org/10.1525/mp.2002.19.3.285>
- Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: Nonhuman primates illuminate human speech processing. *Nature Neuroscience*, 12(6), 718–724. <https://doi.org/10.1038/nn.2331>
- Rauschecker, J. P., & Tian, B. (2000). Mechanisms and streams for processing of “what” and “where” in auditory cortex. *Proceedings of the National Academy of Sciences*, 97(22), 11800–11806. <https://doi.org/10.1073/pnas.97.22.11800>
- Rohrmeier, M., & Koelsch, S. (2012). Predictive information processing in music cognition. A critical review. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*, 83(2), 164–175. <https://doi.org/10.1016/j.ijpsycho.2011.12.010>
- Rolls, E. T., Cheng, W., & Feng, J. (2020). The orbitofrontal cortex: Reward, emotion and depression. *Brain Communications*, 2, 2. <https://doi.org/10.1093/braincomms/fcaa196>
- Rousseau, J.-J. (1781). *Essay on the origin of languages. Translated and edited by John T. Scott*. University Press of New England.
- Sachs, C. (1943). *The rise of music in the ancient world east and west*. W.W. Norton & Company.
- Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: Brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences*, 10(1), 24–30. <https://doi.org/10.1016/j.tics.2005.11.009>
- Schmithorst, V. J. (2005). Separate cortical networks involved in music perception: Preliminary functional MRI evidence for modularity of music processing. *NeuroImage*, 25(2), 444–451. <https://doi.org/10.1016/j.neuroimage.2004.12.006>
- Schoenberg, A. (1922). *A theory of harmony* (3rd ed). Faber And Faber Ltd.
- Schulze, K., Mueller, K., & Koelsch, S. (2011). Neural correlates of strategy use during auditory working memory in musicians and non-musicians. *European Journal of Neuroscience*, 33(1), 189–196. <https://doi.org/10.1111/j.1460-9568.2010.07470.x>
- 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., Ripollés, P., Särkämö, T., Leo, V., Rodríguez-Fornells, A., Saunavaara, J., & Soinila, S. (2017). Tracing the neural basis of music: Deficient structural connectivity underlying acquired amusia. *Cortex*, 97, 255–273. <https://doi.org/10.1016/j.cortex.2017.09.028>
- 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>
- Slevc, L. R., & Okada, B. M. (2015). Processing structure in language and music: A case for shared reliance on cognitive control. *Psychonomic Bulletin & Review*, 22(3), 637–652. <https://doi.org/10.3758/s13423-014-0712-4>
- Spada, D., Verga, L., Iadanza, A., Tettamanti, M., & Perani, D. (2014). The auditory scene: An fMRI study on melody and accompaniment in professional pianists. *NeuroImage*, 102(P2), 764–775. <https://doi.org/10.1016/j.neuroimage.2014.08.036>
- Steele, J. (1775). *An essay towards establishing the melody and measure of speech to be expressed and perpetuated by certain symbols*. Bowyer and Nichols.
- Steinbeis, N., & Koelsch, S. (2008). Shared neural resources between music and language indicate semantic processing of musical tension-resolution patterns. *Cerebral Cortex*, 18(5), 1169–1178. <https://doi.org/10.1093/cercor/bhm149>
- Steinbeis, N., Koelsch, S., & Sloboda, J. A. (2006). The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological, and neural responses. *Journal of Cognitive Neuroscience*, 18(8), 1380–1393. <https://doi.org/10.1162/jocn.2006.18.8.1380>
- Tillmann, B., Janata, P., & Bharucha, J. J. (2003). Activation of the inferior frontal cortex in musical priming. *Annals of the New York Academy of Sciences*, 999, 209–211. <https://doi.org/10.1196/annals.1284.031>
- Tillmann, B., Koelsch, S., Escoffier, N., Bigand, E., Lalitte, P., Friederici, A. D., & von Cramon, D. Y. (2006). Cognitive priming in sung and instrumental music: Activation of inferior frontal cortex. *NeuroImage*, 31(4), 1771–1782. <https://doi.org/10.1016/j.neuroimage.2006.02.028>
- Trost, W., Labbé, C., & Grandjean, D. (2017). Rhythmic entrainment as a musical affect induction mechanism. *Neuropsychologia*, 96, 96–110. <https://doi.org/10.1016/j.neuropsychologia.2017.01.004>
- Tsai, C.-G., & Li, C.-W. (2019). Increased activation in the left ventrolateral prefrontal cortex and temporal pole during tonality change in music. *Neuroscience Letters*, 696, 162–167. <https://doi.org/10.1016/j.neulet.2018.12.019>
- Turkeltaub, P. E., Eden, G. F., Jones, K. M., & Zeffiro, T. A. (2002). Meta-analysis of the functional neuroanatomy of single-word reading: Method and validation. *NeuroImage*, 16(3), 765–780. <https://doi.org/10.1006/nimg.2002.1131>
- Vuust, P., Roepstorff, A., Wallentin, M., Mouridsen, K., & Østergaard, L. (2006). It don’t mean a thing... keeping the rhythm during polyrhythmic tension, activates language areas (BA47). *NeuroImage*, 31(2), 832–841. <https://doi.org/10.1016/j.neuroimage.2005.12.037>
- Vuust, P., Wallentin, M., Mouridsen, K., Østergaard, L., & Roepstorff, A. (2011). Tapping polyrhythms in music activates language areas. *Neuroscience Letters*, 494(3), 211–216. <https://doi.org/10.1016/j.neulet.2011.03.015>

- Vuust, P., & Witek, M. A. G. (2014). Rhythmic complexity and predictive coding: A novel approach to modeling rhythm and meter perception in music. *Frontiers in Psychology, 5*, 1111. <https://doi.org/10.3389/fpsyg.2014.01111>
- Wallaschek, R. (1891). On the origins of music. *Mind: A Quarterly Review of Psychology and Philosophy, 16*, 375–386. <https://doi.org/10.1093/mind/os-XVI.63.375>
- Weiller, C., Reisert, M., Peto, I., Hennig, J., Makris, N., Petrides, M., & Egger, K. (2021). The ventral pathway of the human brain: A continuous association tract system. *NeuroImage, 234*, 117977. <https://doi.org/10.1016/j.neuroimage.2021.117977>
- Wojtasik, M., Bludau, S., Eickhoff, S. B., Mohlberg, H., Gerboga, F., Caspers, S., & Amunts, K. (2020). Cytoarchitectonic characterization and functional decoding of four new areas in the human lateral orbitofrontal cortex. *Frontiers in Neuroanatomy, 14*, 2. <https://doi.org/10.3389/fnana.2020.00002>
- Zatorre, R. J., Belin, & Penhune. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences, 6613*(1994), 37–46. [https://doi.org/10.1016/S1364-6613\(00\)01816-7](https://doi.org/10.1016/S1364-6613(00)01816-7)
- Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: Auditory-motor interactions in music perception and production. *Nature Reviews Neuroscience, 8*, 547–558. <https://doi.org/10.1038/nrn2152>
- Zatorre, R. J., Evans, A. C., & Meyer, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. *Journal of Neuroscience, 14*(4), 1908–1919. <https://doi.org/10.1523/jneurosci.14-04-01908.1994>
- Zatorre, R. J., Evans, A. C., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science (New York, N.Y.), 256*(5058), 846–849. <https://doi.org/10.1126/science.1589767>