frontiers
in Psychology

The relationship between the neural computations for speech and music perception is context-dependent: an activation likelihood estimate study

Arianna N. LaCroix, Alvaro F. Diaz and Corianne Rogalsky *

Communication Neuroimaging and Neuroscience Laboratory, Department of Speech and Hearing Science, Arizona State University, Tempe, AZ, USA

The relationship between the neurobiology of speech and music has been investigated for more than a century. There remains no widespread agreement regarding how (or to what extent) music perception utilizes the neural circuitry that is engaged in speech processing, particularly at the cortical level. Prominent models such as Patel's Shared Syntactic Integration Resource Hypothesis (SSIRH) and Koelsch's neurocognitive model of music perception suggest a high degree of overlap, particularly in the frontal lobe, but also perhaps more distinct representations in the temporal lobe with hemispheric asymmetries. The present meta-analysis study used activation likelihood estimate analyses to identify the brain regions consistently activated for music as compared to speech across the functional neuroimaging (fMRI and PET) literature. Eighty music and 91 speech neuroimaging studies of healthy adult control subjects were analyzed. Peak activations reported in the music and speech studies were divided into four paradigm categories: passive listening, discrimination tasks, error/anomaly detection tasks and memory-related tasks. We then compared activation likelihood estimates within each category for music vs. speech, and each music condition with passive listening. We found that listening to music and to speech preferentially activate distinct temporo-parietal bilateral cortical networks. We also found music and speech to have shared resources in the left pars opercularis but speech-specific resources in the left pars triangularis. The extent to which music recruited speech-activated frontal resources was modulated by task. While there are certainly limitations to meta-analysis techniques particularly regarding sensitivity, this work suggests that the extent of shared resources between speech and music may be task-dependent and highlights the need to consider how task effects may be affecting conclusions regarding the neurobiology of speech and music.

Keywords: music perception, speech perception, fMRI, meta-analysis, Broca's area

Introduction

The relationship between the neurobiology of speech and music has been investigated and debated for nearly a century. (Henschen, 1924; Luria et al., 1965; Frances et al., 1973; Peretz, 2006; Besson et al., 2011). Early evidence from case studies of brain-damaged individuals suggested a dissociation of aphasia and amusia (Yamadori et al., 1977; Basso and Capitani, 1985;

Peretz et al., 1994, 1997; Steinke et al., 1997; Patel et al., 1998b; Tzortzis et al., 2000; Peretz and Hyde, 2003). However, more recent patient work examining specific aspects of speech and music processing indicate at least some overlap in deficits across the two domains. For example, patients with Broca's aphasia have both linguistic and harmonic structure deficits, and patients with amusia exhibit pitch deficits in both speech and music (Patel, 2003, 2005, 2013). Electrophysiological (e.g., ERP) studies also suggest shared resources between speech and music; for example, syntactic and harmonic violations elicit indistinguishable ERP responses such as the P600 response, which is hypothesized to originate from anterior temporal or inferior frontal regions (Patel et al., 1998a; Maillard et al., 2011; Sammler et al., 2011). Music perception also interacts with morphosyntactic representations of speech: the early right anterior negativity (ERAN) ERP component sensitive to chord irregularities interacts with the left anterior negativity's (LAN's) response to morphosyntactic violations or irregularities (Koelsch et al., 2005; Steinbeis and Koelsch, 2008b; Koelsch, 2011).

Several studies of trained musicians and individuals with absolute pitch also suggest an overlap between speech and music as there are carry-over effects of musical training onto speech processing performance (e.g., Oechslin et al., 2010; Elmer et al., 2012; for a review see Besson et al., 2011).

There is a rich literature of electrophysiological and behavioral work regarding the relationship between music and language (for reviews see Besson et al., 2011; Koelsch, 2011; Patel, 2012, 2013; Tillmann, 2012; Slevc and Okada, 2015). This work has provided numerous pieces of evidence of overlap between the neural resources of speech and music, including in the brainstem, auditory cortex and frontal cortical regions (Koelsch, 2011). This high degree of interaction between speech and music coincides with Koelsch et al.'s view that speech and music, and therefore the brain networks supporting them, cannot be separated because of their numerous shared properties, i.e., there is a "music-speech continuum" (Koelsch and Friederici, 2003; Koelsch and Siebel, 2005; Koelsch, 2011). However, evidence from brain-damaged patients suggests that music and speech abilities may dissociate, although there are also reports to the contrary (see above). Patel's (2003, 2008, 2012) Shared Syntactic Integration Resource Hypothesis (SSIRH) is in many ways a remedy to the shared-vs.-distinct debate in the realm of structural/syntactic processing. Stemming in part from the patient and electrophysiological findings, Patel proposes that language and music utilize overlapping cognitive resources but also have unique neural representations. Patel proposes that the shared resources reside in the inferior frontal lobe (i.e., Broca's area) and that distinct processes for speech and music reside in the temporal lobes (Patel, 2003).

The emergence of functional neuroimaging techniques such as fMRI have continued to fuel the debate over the contributions of shared vs. distinct neural resources for speech and music. FMRI lacks the high temporal resolution of electrophysiological methods and can introduce high levels of ambient noise potentially contaminating recorded responses to auditory stimuli. However, the greater spatial resolution of fMRI may provide additional information regarding the neural correlates

of speech and music, and MRI scanner noise can be minimized using sparse sampling scanning protocols and reduced-noise continuous scanning techniques (Peelle et al., 2010). Hundreds of fMRI papers have investigated musical processes, and thousands have investigated the neural substrates of speech. Conversely, to our knowledge and as Slevc and Okada (2015) noted, only a few studies have directly compared activations to hierarchical speech and music (i.e., sentences and melodies) using fMRI (Abrams et al., 2011; Fedorenko et al., 2011; Rogalsky et al., 2011). Findings from these studies conflict with the ERP literature (e.g., Koelsch, 2005; Koelsch et al., 2005) in that the fMRI studies identify distinct neuroanatomy and/or activation response patterns for music and speech processing, although there are notable differences across these studies, particularly relating to the involvement of Broca's area in speech and music.

The differences found across neuroimaging studies regarding the overlap of the neural correlates of speech and music likely arise from the tasks used in each of these studies. For example, Rogalsky et al. used passive listening and found no activation of Broca's area to either speech or music compared to rest. Conversely, Fedorenko et al. used a reading/memory probe task for sentences and an emotional ranking for music and found Broca's area to be preferentially activated by speech but also activated by music compared to rest. There is also evidence that the P600, the ERP component that is sensitive to both speech and music violations, is only present when subjects are actively attending to the stimulus (Besson and Faita, 1995; Brattico et al., 2006; Koelsch, 2011). The inclusion of a task may affect not only the brain regions involved, but also reliability of results: an fMRI study of visual tasks reported that tasks with high attentional loads also had the highest reliability measures compared to passive conditions (Specht et al., 2003). This finding in the visual domain suggests the possibility that greater (within and between) subject variability in passive listening conditions may lead to null effects in group-averaged results.

Given the scarcity of within-subject neuroimaging studies of speech and music, it is particularly critical to examine across-study, between-subjects findings to build a better picture regarding the neurobiology of speech and music. A major barrier in interpreting between-subject neuroimaging results is the variety of paradigms and tasks used to investigate speech and music neural resources. Most scientists studying the neurobiology of speech and/or music would likely agree that they are interested in understanding the neural computations employed in naturalistic situations that are driven by the input of speech or music, and the differences between the two. However, explicit tasks such as discrimination or error detection are often used to drive brain responses in part by increasing the subject's attention to the stimuli and/or particular aspects of the stimuli. This may be problematic: the influence of task demands on the functional neuroanatomy recruited by speech is well documented (e.g., Baker et al., 1981; Noesselt et al., 2003; Scheich et al., 2007; Geiser et al., 2008; Rogalsky and Hickok, 2009) and both speech and music processing engage domain-general cognitive, memory, and motor networks in likely distinct, but overlapping ways (Besson et al., 2011). Task effects are known to alter inter and intra hemisphere activations to speech

(Noesselt et al., 2003; Tervaniemi and Hugdahl, 2003; Scheich et al., 2007; Geiser et al., 2008; Rogalsky and Hickok, 2009). For example, there is evidence that right hemisphere fronto-temporal-parietal networks are significantly activated during an explicit task (rhythm judgment) with speech stimuli but not during passive listening to the same stimuli (Geiser et al., 2008). The neurobiology of speech perception, and auditory processing more generally, also can vary based on the type of explicit task even when the same stimuli are used across tasks (Platel et al., 1997; Ni et al., 2000; Von Kriegstein et al., 2003; Geiser et al., 2008; Rogalsky and Hickok, 2009). This phenomenon is also well documented in the visual domain (Corbetta et al., 1990; Chawla et al., 1999; Cant and Goodale, 2007). For example, in the speech domain, syllable discrimination and single-word comprehension performance (as measured by a word-picture matching task) doubly dissociate in stroke patients with aphasia (Baker et al., 1981). Syllable discrimination implicates left-lateralized dorsal frontal-parietal networks, while speech comprehension and passive listening tasks engage mostly mid and posterior temporal regions (Dronkers et al., 2004; Schwartz et al., 2012; Rogalsky et al., 2015). Similarly, contextual effects have been reported regarding pitch: when pitch is needed for linguistic processing, such as in a tonal language, there is a left hemisphere auditory cortex bias, while pitch processing in a melody discrimination task yields a right hemisphere bias (Zatorre and Gandour, 2008). Another example of the importance of context in pitch processing is in vowel perception: vowels and tones have similar acoustic features and when presented in isolation (i.e., just a vowel, not in a consonant-vowel (CV) pair as would typically be perceived in everyday life) no significant differences have been found in temporal lobe activations (Jäncke et al., 2002). However, there is greater superior temporal activation for CVs than tones suggesting that the context of the vowel modulates the temporal networks activated (Jäncke et al., 2002).

One way to reduce the influence of a particular paradigm or task is to use meta-analysis techniques to identify areas of activation that consistently activate to a particular stimulus (e.g., speech, music) across a range of tasks and paradigms. Besson and Schön (2001) noted that meta-analyses of neuroimaging data would provide critical insight into the relationship between the neurobiology of language and music. They also suggested that meta-analyses of music-related neuroimaging data were not feasible due to the sparse number of relevant studies. Now, almost 15 years later, there is a large enough corpus of neuroimaging work to conduct quantitative meta-analyses of music processing with sufficient power. In fact, such meta-analyses have begun to emerge, for specific aspects of musical processing, in relation to specific cognitive functions [e.g., Slevc and Okada's (2015) cognitive control meta-analysis in relation to pitch and harmonic ambiguity], in addition to extensive qualitative reviews (e.g., Tervaniemi, 2001; Jäncke, 2008; Besson et al., 2011; Grahn, 2012; Slevc, 2012; Tillmann, 2012).

The present meta-analysis addresses the following outstanding questions: (1) has functional neuroimaging identified significant distinctions between the functional neuroanatomy of speech and music and (2) how do specific types of tasks affect how music recruits speech-processing networks? We then discuss the implications of our findings for future investigations of the neural computations of language and music.

Materials and Methods

An exhaustive literature search was conducted via Google Scholar to locate published fMRI and PET studies reporting activations to musical stimuli. The following search terms were used to locate papers about music: "fMRI music," "fMRI and music," "fMRI pitch," and "fMRI rhythm." To the best of our knowledge, all relevant journal research articles have been collected for the purposes of this meta-analysis.

All journal articles that became part of the meta-analysis reported peak coordinates for relevant contrasts. Peak coordinates reported in the papers identified by the searches were divided into four categories that encompassed the vast majority of paradigms used in the articles: music passive listening, music discrimination, music error detection, and music memory[1]. Passive listening studies included papers in which participants listened to instrumental melodies or tone sequences with no explicit task as well as studies that asked participants to press a button when the stimulus concluded. Music discrimination studies included those that asked participants to compare two musical stimuli (e.g., related/unrelated, same/different). Music error detection studies included studies that instructed participants to identify a dissonant melody, unexpected note or deviant instrument. The music memory category included papers that asked participants to complete an n-back task, familiarity judgment, or rehearsal (covert or overt) of a melodic stimulus.

Only coordinates from healthy adult, non-musician, control subjects were included. In studies that included a patient group and a control group, only the control group's coordinates were included. Studies were excluded from the final activation likelihood estimate (ALE) if the data did not meet the requirements for being included in ALE calculations, including for the following reasons: coordinates not reported, only approximate anatomical location reported, stereotaxic space not reported, inappropriate contrasts (e.g., speech > music only), activations corresponding to participant's emotional reactions to music, studies of professional/trained musicians, and studies of children.

In addition to collecting the music-related coordinates via an exhaustive search, we also gathered a representative sample of fMRI and PET studies that reported coordinates for passive listening to intelligible speech compared to some type of non-speech control (e.g., tones, noise, rest, visual stimuli).

---

[1]The music categories included studies with stimuli of the following types: instrumental unfamiliar and familiar melodies, tone sequences and individual tones. In comparison, the speech categories described below included studies with stimuli such as individual phonemes, vowels, syllables, words, pseudowords, sentences, and pseudoword sentences. For the purposes of the present study, we have generated two distinct groups of stimuli to compare. However, music and speech are often conceptualized as being two ends of continuum with substantial gray area between the two extremes (Koelsch, 2011). For example, naturally spoken sentences contain rhythmic and pitch-related prosodic features and a familiar melody likely automatically elicits a mental representation of the song's lyrics.

Coordinates corresponding to the following tasks were also extracted: speech discrimination, speech detection, and speech memory. The purpose of these speech conditions is to act as comparison groups for the music groups. Coordinates for this purpose were extracted from six sources: five well-cited review papers, Price (2010), Zheng et al. (2010), Turkeltaub and Coslett (2010), Rogalsky et al. (2011), and Adank (2012) and the brain imaging meta-analysis database Neurosynth.org. The Price (2010), Zheng et al. (2010), Turkeltaub and Coslett (2010), Rogalsky et al. (2011), and Adank (2012) papers yielded a total of 42 studies that fit the aforementioned criteria. An additional 49 relevant papers were found using the Neurosynth.org database with the search criteria "speech perception," "speech processing," "speech," and "auditory working memory." These methods resulted in 91 studies in which control subjects passively listened to speech or completed an auditory verbal memory, speech discrimination, or speech detection task. The passive listening speech condition included studies in which participants listened to speech stimuli with no explicit task as well as studies that asked participants to press a button when the stimulus concluded. Papers were included in the speech discrimination category if they asked participants to compare two speech stimuli (e.g., a same/different task). The speech detection category contained papers that asked participants to detect semantic, intelligibility, or grammatical properties or detect phonological, semantic, or syntactic errors. Studies included in the speech memory category were papers that instructed participants to complete an n-back task or rehearsal (covert or overt) of a speech (auditory verbal) stimulus.

Analyses were conducted using the meta-analysis software GingerALE to calculate ALEs for each condition based on the coordinates collected (Eickhoff et al., 2009, 2012; Turkeltaub et al., 2012). All results are reported in Talairach space. Coordinates originally reported in MNI space were transformed to Talairach space using GingerALE's stereotaxic coordinate converter. Once all coordinates were in Talairach space, each condition was analyzed individually using the following GingerALE parameters: less conservative (larger) mask size, Turkeltaub nonadditive ALE method (Turkeltaub et al., 2012), subject-based FWHM (Eickhoff et al., 2009), corrected threshold of $p < 0.05$ using false discovery rate (FDR), and a minimum cluster volume of $200\,mm^3$. We obtained subtraction contrasts between two given conditions by directly comparing activations between two conditions. To correct for multiple comparisons, each contrast's threshold was set to $p < 0.05$, whole-brain corrected following the FDR algorithm with p value permutations set at 10,000, and a minimum cluster size of $200\,mm^3$ (Eickhoff et al., 2009). ALE statistical maps were rendered onto the Colin Talairach template brain using the software MRIcron (Rorden and Brett, 2000).

Results

Search Results

The literature search yielded 80 music studies (76 fMRI studies, 4 PET studies) and 91 relevant speech papers (88 fMRI, 3 PET studies) meeting the inclusion criteria described above. **Table 1** indicates the number of studies, subjects, and coordinates in each of the four music conditions, as well as for each of the four speech conditions.

TABLE 1 | Activations included in the present meta-analysis.

Condition	Number of studies	Number of subjects	Number of coordinates
Music passive listening	41	540	526
Music discrimination	12	211	168
Music error detection	25	355	489
Music memory	14	190	207
Speech passive listening	31	454	337
Speech discrimination	31	405	318
Speech detection	17	317	248
Speech memory	19	259	324

Passive Listening To Music vs. Passive Listening To Speech

The music passive listening ALE identified large swaths of voxels bilaterally, spanning the length of the superior temporal gyri (STG), as well as additional smaller clusters, including in the bilateral inferior frontal gyrus (pars opercularis), bilateral postcentral gyrus, bilateral insula, left inferior parietal lobule, left medial frontal gyrus, right precentral gyrus, and right middle frontal gyrus (**Figure 1A**, **Table 2**). The speech passive listening ALE also identified bilateral superior temporal regions as well as bilateral precentral and inferior frontal (pars opercularis) regions. Notably, the speech ALE identified bilateral anterior STG, bilateral superior temporal sulcus (i.e., both banks, the middle and superior temporal gyri) and left inferior frontal gyrus (pars triangularis) regions not identified by the music ALE (**Figure 1A**, **Table 2**). ALEs used a threshold of $p < 0.05$, FDR corrected.

Pairwise contrasts of passive listening to music vs. passive listening to speech were calculated to identify any brain regions that were significantly activated more by speech or music, respectively. Results were as follows ($p < 0.05$, FDR corrected): the speech > music contrast identified significant regions on both banks of the bilateral superior temporal sulcus extending the length of the left temporal lobe and mid/anterior right temporal lobe, left inferior frontal lobe (pars triangularis), left precentral gyrus, and left postcentral gyrus regions. Music > speech identified bilateral insula and bilateral superior temporal/parietal operculum clusters as well as a right inferior frontal gyrus region (**Figure 1B**, **Table 2**). These results coincide with previous reports of listening to speech activating a lateral temporal network particularly in the superior temporal sulcus and extending into the anterior temporal lobe, while listening to music activated a more dorsal medial temporal-parietal network (Jäncke et al., 2002; Rogalsky et al., 2011). These results also coincide with Fedorenko et al.'s (2011) finding that Broca's area, the pars triangularis in particular, is preferentially responsive to language stimuli.
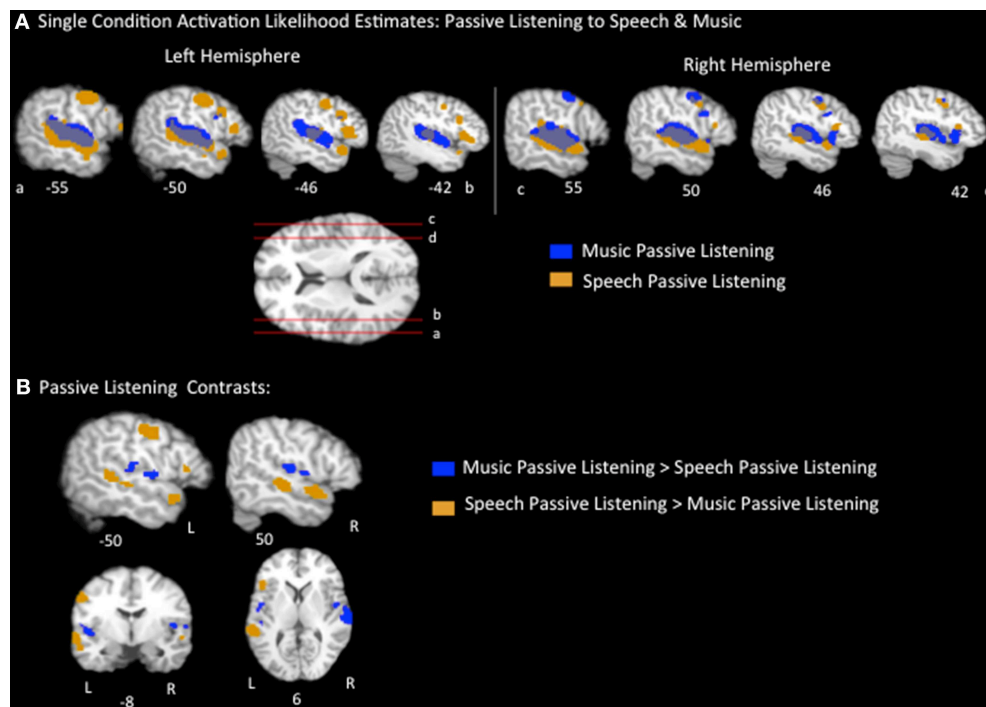
FIGURE 1 | (A) Representative sagittal slices of the ALE for passive listening to speech, $p < 0.05$, corrected, overlaid on top of the passive music listening ALE. **(B)** Speech vs. music passive listening contrasts results, $p < 0.05$ corrected.

Music Tasks vs. Speech Tasks

The passive listening ALE results identify distinct and overlapping regions of speech and music processing. We now turn to the question of how do these distinctions change as a function of the type of task employed? First, ALEs were computed for each music task condition, $p < 0.05$ FDR corrected (**Figure 1**, **Table 2**). The music task conditions' ALEs all significantly identified bilateral STG and bilateral precentral gyrus, and inferior parietal regions, overlapping with the passive listening music ALE (**Figure 2**). The tasks also activated additional inferior frontal and inferior parietal regions not identified by the passive listening music ALE; these differences are discussed in a subsequent section.

To compare the brain regions activated by each music task to those activated by speech in similar tasks, pairwise contrasts of the ALEs for each music task vs. its corresponding speech task group were calculated (**Figure 3**, **Table 2**). Music discrimination > speech discrimination identified regions including bilateral inferior frontal gyri (pars opercularis), bilateral pre and postcentral gyri, bilateral medial frontal gyri, left inferior parietal lobule, and left cerebellum, whereas speech discrimination > music discrimination identified bilateral regions in the anterior superior temporal sulci (including both superior and middle temporal gyri). Music detection > speech detection identified a bilateral group of clusters spanning the superior temporal gyri, bilateral precentral gyri, bilateral insula and bilateral inferior parietal regions, as well as clusters in the right middle frontal gyrus. Speech detection > music detection

identified bilateral superior temporal sulci regions as well as left inferior frontal regions (pars triangularis and pars opercularis). Music memory > speech memory identified a left posterior superior temporal/inferior parietal region and bilateral medial frontal regions; speech memory > music memory identified left inferior frontal gyrus (pars opercularis and pars triangularis) and bilateral superior and middle temporal gyri.

In sum, the task pairwise contrasts in many ways mirror the passive listening contrast: music tasks activated more dorsal/medial superior temporal and inferior parietal regions, while speech tasks activated superior temporal sulcus regions, particularly in the anterior temporal lobe. In addition, notable differences were found in Broca's area and its right hemisphere homolog: in discrimination tasks music significantly activated Broca's area (specifically the pars opercularis) more than speech. However, in detection and memory tasks speech activated Broca's area (pars opercularis and pars triangularis) more than music. The right inferior frontal gyrus responded equally to speech and music in both detection and memory tasks, but responded more to music than speech in discrimination tasks. Also notably, in the memory tasks, music activated a lateral superior temporal/inferior parietal cluster (in the vicinity of Hickok and Poeppel's "area Spt") more than speech while an inferior frontal cluster including the pars opercularis was activated more for speech than music. Both area Spt and the pars opercularis previously have been implicated in a variety of auditory working memory tasks (including speech and pitch working memory) in both lesion patients and control subjects (Koelsch and Siebel,

TABLE 2 | Locations, peaks and cluster size for significant voxel clusters for each condition's ALE and for each contrast of interest.

Condition	Anatomical locations	Peak coordinates	Voxels
Music passive listening	Left inferior frontal gyrus (pars opercularis)*	−46, 10, 26	32
	Left medial frontal gyrus*, left subcallosal gyrus	−2, 26,−14	65
	Left medial frontal gyrus*	−2, 2, 62	48
	Left postcentral gyrus*, left inferior parietal lobule	−34,−36, 54	27
	Left superior temporal gyrus*, left transverse temporal gyrus, left middle temporal gyrus, left insula	−52,−20, 6	2073
	Right inferior frontal gyrus*	48, 10, 28	43
	Right precentral gyrus*, right postcentral gyrus, right middle frontal gyrus	52,−2, 44	173
	Right superior temporal gyrus*, right transverse temporal gyrus, right middle temporal gyrus, right insula	58,−20, 6	2154
	Right insula*, right inferior frontal gyrus, right precentral gyrus	42, 14, 0	206
	Right lingual gyrus*, right culmen	16,−54,−2	27
Music discrimination	Left medial frontal gyrus*, left middle frontal gyrus	−8,−4, 58	224
	Left precentral gyrus*, left postcentral gyrus, left inferior parietal lobule	−48,−12, 48	259
	Left precentral gyrus*, left inferior frontal gyrus (pars opercularis)	−50, 2, 26	67
	Left superior temporal gyrus*, left transverse temporal gyrus, left precentral gyrus	−54,−16, 8	239
	Left superior temporal gyrus*, left middle temporal gyrus	−58,−34, 8	92
	Left insula*, left inferior frontal gyrus (pars triangularis)	−34, 22, 2	48
	Left cerebellum*	−28,−62,−24	127
	Right inferior frontal gyrus*, right middle frontal gyrus	52, 12, 28	58
	Right precentral gyrus*, right middle frontal gyrus	46,−6, 44	170
	Right superior temporal gyrus*, right middle temporal gyrus	62,−24, 8	310
	Right superior temporal gyrus*, right precentral gyrus, right insula	50, 6,−2	91
Music error detection	Left medial frontal gyrus*	−4,−4, 58	49
	Left superior temporal gyrus*, Let transverse temporal gyrus, Left postcentral gyrus, left insula	−50,−18, 8	1448
	Left inferior parietal lobule*, left supramarginal gyrus, left angular gyrus	−40,−48, 40	41
	Left lentiform nucleus*, left putamen	−22, 6, 10	263
	Right middle frontal gyrus*	36, 42, 18	43
	Right middle frontal gyrus*, right precentral gyrus	32,−4, 56	35
	Right superior frontal gyrus*, right medial frontal gyrus, left superior frontal gyrus, left medial frontal gyrus	2, 10, 52	95
	Right superior temporal gyrus*, right transverse temporal gyrus, right insula, right precentral gyrus, right middle temporal gyrus, right claustrum	50,−18, 6	1228
	Right parahippocampal gyrus*	22,−14,−12	36
	Right inferior parietal lobule*, right supramarginal gyrus	36,−44, 40	103
	Right insula*, right inferior frontal gyrus	32, 22, 12	329
	Right lentiform nucleus*, right putamen, right caudate	18, 6, 12	144
	Right thalamus*	12,−16, 8	33
	Right cerebellum*	26,−50,−26	28
Music memory	Left inferior frontal gyrus (pars opercularis)*, left precentral gyrus, left middle frontal gyrus	−50, 4, 26	206
	Left inferior frontal gyrus (pars triangularis*, pars orbitalis), left insula	−34, 24,−2	57
	Left inferior frontal gyrus (pars triangularis)*	−44, 26, 10	25
	Left medial frontal gyrus*	−4, 52, 12	31
	Left middle frontal gyrus*	−32, 4, 54	29
	Left precentral gyrus*	−44,−10, 42	33
	Left superior frontal gyrus*, left medial frontal gyrus, right superior frontal gyrus, right medial frontal gyrus	−0, 12, 50	373
	Left middle temporal gyrus*	−50,−20,−10	72
	Left middle temporal gyrus*, left superior temporal gyrus	−46, 4,−18	35

(Continued)

TABLE 2 | Continued

Condition	Anatomical locations	Peak coordinates	Voxels
	Left inferior parietal lobule*, left superior temporal gyrus, left middle temporal gyrus, left supramarginal gyrus	−48,−44, 22	224
	Left thalamus*	−14,−14, 14	37
	Right inferior frontal gyrus*, right insula, right claustrum	32, 26, 8	90
	Right middle frontal gyrus*	38, 44, 14	27
	Right superior temporal gyrus*, right middle temporal gyrus	54,−38, 10	35
	Right parahippocampal gyrus*, right hippocampus	30,−10,−20	35
	Right cerebellum*	30,−56,−18	47
Speech passive listening	Left inferior frontal gyrus (pars triangularis, pars opercularis)*, left insula, left precentral gyrus	−44, 20, 8	296
	Left inferior frontal gyrus (pars triangularis*, pars opercularis), left middle frontal gyrus, left precentral gyrus	−48, 10, 28	162
	Left precentral gyrus*, left postcentral gyrus	−52,−10, 40	294
	Left medial frontal gyrus*, left superior frontal gyrus, left medial frontal gyrus, left cingulate gyrus	−8, 8, 50	164
	Left superior temporal gyrus*, Left middle temporal gyrus, left postcentral gyrus, left transverse temporal gyrus	−58,−14,−2	2101
	Left superior temporal gyrus*	−46, 12,−14	107
	Left fusiform gyrus*, left inferior occipital gyrus, left middle occipital gyrus	−38,−78,−12	35
	Right inferior frontal gyrus*, right insula, right precentral gyrus	44, 18, 10	81
	Right middle frontal gyrus*, right precentral gyrus, right inferior frontal gyrus	46, 2, 38	118
	Right superior temporal gyrus*, right middle temporal gyrus, Right insula, right precentral gyrus, right transverse temporal gyrus	−52, 20, 0	1800
Speech discrimination	Left inferior frontal gyrus (pars orbitalis*, pars triangularis), left insula, left middle frontal gyrus	−38, 26,−4	115
	Left inferior frontal gyrus (pars triangularis*, pars opercularis), left precentral gyrus	−44, 20, 10	44
	Left middle frontal gyrus*, left inferior frontal gyrus (pars triangularis, pars opercularis)	−46, 16, 30	187
	Left middle frontal gyrus*, left precentral gyrus	−46,−0, 42	26
	Left superior temporal gyrus*, left postcentral gyrus, left transverse temporal gyrus, left middle temporal gyrus	−58,−20, 4	1737
	Left thalamus*, left caudate	−14,−16, 10	147
	Left cerebellum	−38,−60,−16	36
	Right inferior frontal gyrus*, right precentral gyrus, right insula	46, 20, 4	38
	Right superior temporal gyrus*, right middle temporal gyrus, right transverse temporal gyrus, right insula	58,−14, 0	1223
	Right precuneus*, right cuneus	4,−78, 38	34
Speech detection	Left inferior frontal gyrus (pars opercularis)*, left middle frontal gyrus, left insula	−48, 10, 22	361
	Left inferior frontal gyrus (pars triangularis)*	−48, 28, 12	101
	Left inferior frontal gyrus (pars triangularis*, pars orbitalis), left insula	−34, 24, 2	61
	Left postcentral gyrus*, left precentral gyrus	−50,−12, 46	92
	Left medial frontal gyrus*, left superior frontal gyrus	−6,−6, 60	54
	Left superior temporal gyrus*, left middle temporal gyrus, left transverse temporal gyrus	−60,−22,−2	1010
	Left superior temporal gyrus*, left supramarginal gyrus, left inferior parietal lobule	−60,−42, 20	66
	Left superior temporal gyrus*, left middle temporal gyrus	−50, 12,−14	34
	Left superior temporal gyrus*	−42, 18,−24	28
	Left transverse temporal gyrus*, left superior temporal gyrus	−36,−30, 12	38
	Left precuneus*, left superior parietal lobule, left inferior parietal lobule	−30,−62, 40	66
	Right inferior frontal gyrus*, right insula	34, 24, 6	62
	Right inferior frontal gyrus*	40, 24,−4	31
	Right inferior frontal gyrus*	52, 8, 22	29
	Right superior temporal gyrus*, right transverse temporal gyrus, right middle temporal gyrus	58,−14, 4	788
	Right middle frontal gyrus*	48, 14, 32	36

(Continued)

TABLE 2 | Continued

| Condition	Anatomical locations	Peak coordinates	Voxels
Speech memory	Left middle frontal gyrus*, left inferior frontal gyrus (pars triangularis, pars opercularis), left precentral gyrus	−50, 22, 22	476
	Left superior frontal gyrus*, left medial frontal gyrus, right medial frontal gyrus, right superior frontal gyrus	−2, 4, 56	73
	Left precentral gyrus*, left postcentral gyrus	−50,−10, 44	127
	Left insula*, left inferior frontal gyrus (pars triangularis), left claustrum	−30, 18, 4	39
	Left superior temporal gyrus*, left middle temporal gyrus, left insula	−62,−24, 6	937
	Left superior temporal gyrus*, left middle temporal gyrus	−50, 10,−10	62
	Left superior parietal lobule*, left precuneus, left inferior parietal lobule	−30,−62, 48	109
	Left inferior parietal lobule*	−40,−46, 50	93
	Left caudate*, left thalamus	−16,−2, 16	36
	Left cerebellum*, left fusiform gyrus	−40,−44,−20	67
	Right superior temporal gyrus*, right middle temporal gyrus, right transverse temporal gyrus	58,−14, 0	773
	Right superior temporal gyrus*, right middle temporal gyrus	48, 8,−14	58
	Right cerebellum	24,−64,−16	50
Music passive > speech passive	Left insula*, left superior temporal gyrus	−44,−6, 2	148
	Left superior temporal gyrus*, left insula, left middle temporal gyrus	−42,−40, 14	146
	Left subcallosal gyrus*, left medial frontal gyrus, left anterior cingulate	−4, 22,−14	53
	Right inferior frontal gyrus*, right insula	44, 18,−2	49
	Right superior temporal gyrus*, right postcentral gyrus, right transverse temporal gyrus, right precentral gyrus, right insula	66,−20, 10	457
Music passive < speech passive	Left inferior frontal gyrus (pars triangularis)*, left precentral gyrus	−42, 30, 2	177
	Left precentral gyrus*, left postcentral gyrus	−56,−10, 40	191
	Left middle temporal gyrus*, left inferior temporal gyrus, left superior temporal gyrus	−56,−12,−12	856
	Left middle temporal gyrus*, left superior temporal gyrus	−50, 6,−18	91
	Left cingulate gyrus*, left medial frontal gyrus, left superior frontal gyrus	−10, 4, 46	70
	Right middle temporal gyrus*, right superior temporal gyrus, right insula	56,−22,−8	277
	Right middle temporal gyrus*, right superior temporal gyrus	52, 2,−12	167
Music discrimination > speech discrimination	Left inferior frontal gyrus (pars opercularis)*, left precentral gyrus	−52, 4, 24	56
	Left postcentral gyrus*, left inferior parietal lobule, left precentral gyrus	−48,−18, 44	253
	Left medial frontal gyrus*, left superior frontal gyrus, right medial frontal gyrus, right superior frontal gyrus	−8,−6, 54	224
	Left superior temporal gyrus*, left transverse temporal gyrus, left precentral gyrus	−52,−10, 8	122
	Left cerebellum	−28,−64,−28	114
	Right inferior frontal gyrus*, right middle frontal gyrus	50, 8, 26	53
	Right precentral gyrus*, right middle frontal gyrus	36,−6, 42	170
	Right precentral gyrus*, right insula, right superior temporal gyrus	48, 4, 8	91
	Right superior temporal gyrus*, right transverse temporal gyrus	66,−26, 10	93
Music discrimination < speech discrimination	Left middle temporal gyrus*, left superior temporal gyrus	−62,−18,−8	456
	Right middle temporal gyrus*, right superior temporal gyrus	66,−8,−4	38
Music detection > speech detection	Left insula*, left superior temporal gyrus, left precentral gyrus	−40,−16, 8	126
	Left insula*, left superior temporal gyrus	−42, 4,−6	76
	Left superior temporal gyrus*, left transverse temporal gyrus	−48,−34, 16	131
	Right insula*, right transverse temporal gyrus, right superior temporal gyrus	44,−10,−4	507
	Right middle frontal gyrus*, right insula	38, 16, 24	78
	Right middle frontal gyrus*, right precentral gyrus	32,−4, 54	35
Music detection < speech detection	Left inferior frontal gyrus (par opercularis)*	−56, 16, 20	240

(Continued)

TABLE 2 | Continued

| Condition	Anatomical locations	Peak coordinates	Voxels
	Left inferior frontal gyrus (pars triangularis)*	−52, 28, 12	101
	Left middle temporal gyrus*, left superior temporal gyrus, left transverse temporal gyrus	−60,−32,−2	561
	Left superior temporal gyrus*	−44, 18,−24	28
	Right middle temporal gyrus*, right superior temporal gyrus	62,−12,−4	361
Music memory > speech memory	Left cingulate gyrus*, left superior frontal gyrus, left medial frontal gyrus, right cingulate gyrus	−6, 20, 32	161
	Left superior temporal gyrus*, left supramarginal gyrus, left inferior parietal lobule	−46,−48, 14	45
Music memory < speech memory	Left inferior frontal gyrus (pars triangularis*, pars opercularis)	−54, 24, 20	80
	Left superior temporal gyrus*, left middle temporal gyrus	−60,−16, 6	606
	Right superior temporal gyrus*, right middle temporal gyrus, right transverse temporal gyrus	52,−26, 2	506
Music passive listening > music discrimination	Left insula*, left superior temporal gyrus	−42,−12,−4	116
	Left superior temporal gyrus*, left insula	−42,−42, 12	261
	Right superior temporal gyrus*, right insula	52,−12, 4	157
Music passive listening < music discrimination	Left medial frontal gyrus*, right medial frontal gyrus	−8,−6, 54	165
	Left precentral gyrus*, left superior temporal gyrus	−52, 2, 8	80
	Left postcentral gyrus*, left inferior parietal lobule, left precentral gyrus	−46,−18, 46	228
	Left cerebellum*	−24,−62,−24	90
	Right precentral gyrus*, right middle frontal gyrus	44,−6, 42	105
	Right precentral gyrus*, right insula, right superior temporal gyrus	50, 6, 6	30
Music passive > music error detection	Left middle temporal gyrus*	−58,−32, 0	82
	Left superior temporal gyrus*	−58,−10, 4	81
	Right precentral gyrus*, right middle frontal gyrus	50, 2, 48	64
	Right postcentral gyrus*, right superior temporal gyrus	62,−24, 16	44
	Right superior temporal gyrus*, right transverse temporal gyrus, right middle temporal gyrus, right precentral gyrus	60,−16, 0	336
Music passive < music error detection	Left medial frontal gyrus*	−4,−8, 56	30
	Left superior frontal gyrus*, left medial frontal gyrus, right superior frontal gyrus	−0, 8, 48	93
	Left postcentral gyrus*, left transverse temporal gyrus, left precentral gyrus	−52,−22, 16	79
	Left inferior parietal lobule*, left supramarginal gyrus	−40,−48, 38	37
	Left superior temporal gyrus*, left precentral gyrus	−52, 2, 4	92
	Left insula*, left superior temporal gyrus, left transverse temporal gyrus	−40,−28, 14	67
	Left lentiform nucleus*, left caudate	−18, 10, 8	211
	Right inferior parietal lobule*, right supramarginal gyrus	36,−50, 42	101
	Right insula*, right inferior frontal gyrus, right middle frontal gyrus	38, 18, 16	227
	Right insula*, right superior temporal gyrus	40,−20, 16	139
	Right insula*, right superior temporal gyrus	42,−8, 0	42
	Right caudate*, right lentiform nucleus	14, 6, 14	143
	Right thalamus*	14,−18, 6	32
	Right cerebellum*	28,−54,−26	28
Music passive listening > music memory	Left superior temporal gyrus*, left middle temporal gyrus, left insula	−54,−22, 6	943
	Right superior temporal gyrus*, right insula, right postcentral gyrus, right precentral gyrus right transverse temporal gyrus, right middle temporal gyrus	52,−20, 4	1350
	Right insula*, right inferior frontal gyrus, right superior temporal gyrus	46, 10, 2	32
Music passive listening < music memory	Left inferior frontal gyrus (pars opercularis)*, left precentral gyrus	−44, 4, 30	79
	Left inferior frontal gyrus (pars orbitalis)*, left insula	−32, 24,−6	53
	Left middle frontal gyrus*, left inferior frontal gyrus (pars triangularis)	−42, 18, 28	29
	Left middle frontal gyrus*	−32, 6, 54	29

(Continued)

TABLE 2 | Continued

Condition	Anatomical locations	Peak coordinates	Voxels
	Left superior frontal gyrus*, left medial frontal gyrus, right medial frontal gyrus, right superior frontal gyrus	−0, 8, 48	329
	Left superior temporal gyrus*, left middle temporal gyrus	−44,−20,−10	69
	Left inferior parietal lobule*, left supramarginal gyrus, left superior temporal gyrus	−54,−44, 28	89
	Left thalamus*	−10,−16, 14	37
	Right inferior frontal gyrus*, right insula, right claustrum	32, 26, 4	83
	Right parahippocampal gyrus*, right hippocampus	32,−12,−24	35
Speech passive listening > music discrimination	Left middle temporal gyrus*, left superior temporal gyrus	−56,−20,−8	298
	Right middle temporal gyrus*, right superior temporal gyrus	56,−18,−4	308
Speech passive listening < music discrimination	Left precentral gyrus*, left superior temporal gyrus	−52, 2, 8	105
	Left postcentral gyrus*, left precentral gyrus, left inferior parietal lobule	−50,−14, 52	199
	Left cerebellum	−28,−64,−28	127
	Right inferior frontal gyrus*, right middle frontal gyrus	50, 10, 30	50
	Right medial frontal gyrus*, right superior frontal gyrus, left medial frontal gyrus, left superior frontal gyrus	2,−6, 62	166
	Right precentral gyrus*, right middle frontal gyrus	38,−8, 42	67
	Right superior temporal gyrus*	64,−26, 8	76
	Right superior temporal gyrus*, right precentral gyrus, right superior temporal gyrus	50, 6, 4	47
Speech passive listening > music detection	Left inferior frontal gyrus (pars triangularis*, pars opercularis),	−50, 22, 10	107
	Left middle frontal gyrus*, left precentral gyrus, left postcentral gyrus,	−54, 2, 40	138
	Left middle temporal gyrus*, left superior temporal gyrus, left inferior temporal gyrus	−60,−8,−10	1052
	Left superior temporal gyrus*	−48, 16,−16	29
	Right middle temporal gyrus*, right superior temporal gyrus	60,−18,−8	651
Speech passive listening < music detection	Left precentral gyrus*, left superior temporal gyrus	−52, 2, 6	54
	Left insula*, left superior temporal gyrus	−50,−20, 16	430
	Left insula*, left superior temporal gyrus	−40, 6,−4	31
	Left inferior parietal lobule*, left supramarginal gyrus	−42,−50, 38	40
	Left lentiform nucleus*, left claustrum, left insula	−20, 6, 6	203
	Right middle frontal gyrus*, right inferior frontal gyrus, right insula	42, 16, 32	220
	Right middle frontal gyrus*, right precentral gyrus	30,−6, 56	35
	Right superior frontal gyrus*, right middle frontal gyrus	32, 44, 16	40
	Right superior frontal gyrus*	4, 12, 54	36
	Right insula*, right transverse temporal gyrus, right superior temporal gyrus, right precentral gyrus	44,−12, 12	519
	Right inferior parietal lobule*, right supramarginal gyrus	40,−48, 38	103
	Right thalamus*, right caudate	8,−2, 10	142
	Right thalamus*	10,−14, 6	33
Speech passive listening > music memory [-10pt]	Left middle temporal gyrus*, left middle temporal gyrus, left transverse temporal gyrus	−58,−38,−4	1256
	Right superior temporal gyrus*, right transverse temporal gyrus, right middle temporal gyrus, right postcentral gyrus	58,−2, 2	1056
Speech passive listening < music memory [-10pt]	Left inferior frontal gyrus (pars orbitalis*, pars triangularis)	−32, 24,−6	31
	Left medial frontal gyrus*, left superior frontal gyrus, left cingulate, right superior frontal gyrus, right medial frontal gyrus, right cingulate	−0, 22, 46	336
	Left precentral gyrus*, left inferior frontal gyrus (pars opercularis)	−52, 2, 30	34
	Left precentral gyrus*	−38,−10, 38	32
	Left supramarginal gyrus*, left inferior parietal lobule, left superior temporal gyrus	−46,−46, 26	113
	Left inferior parietal lobule*, left postcentral gyrus	−48,−36, 48	71
	Left middle temporal gyrus*, left superior temporal gyrus	−46,−24,−10	44
	Right insula*, right inferior frontal gyrus, right claustrum	34, 22, 4	48

The x, y, z coordinates are in Talairach space and refer to the peak voxel activated in each contrast. All contrasts are thresholded at p = 0.05. Asterisks indicate anatomical location of peak voxel.
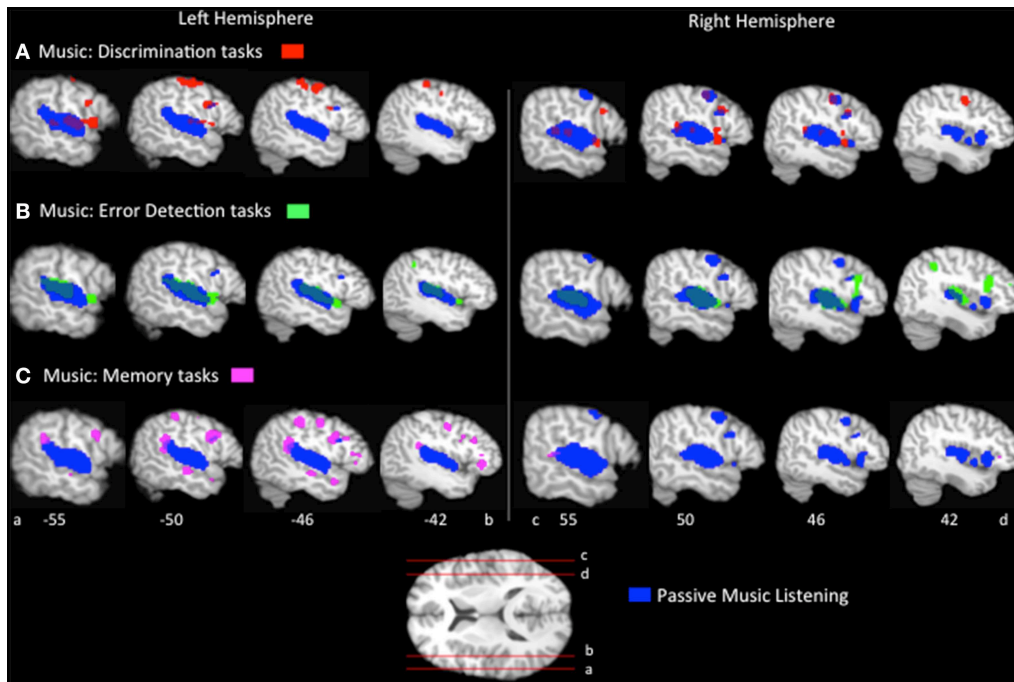
FIGURE 2 | Representative sagittal slices of the ALEs for the (A) music discrimination, (B) music error detection and (C) music memory task conditions, $p < 0.05$, corrected, overlaid on top of the passive music listening ALE for comparison.
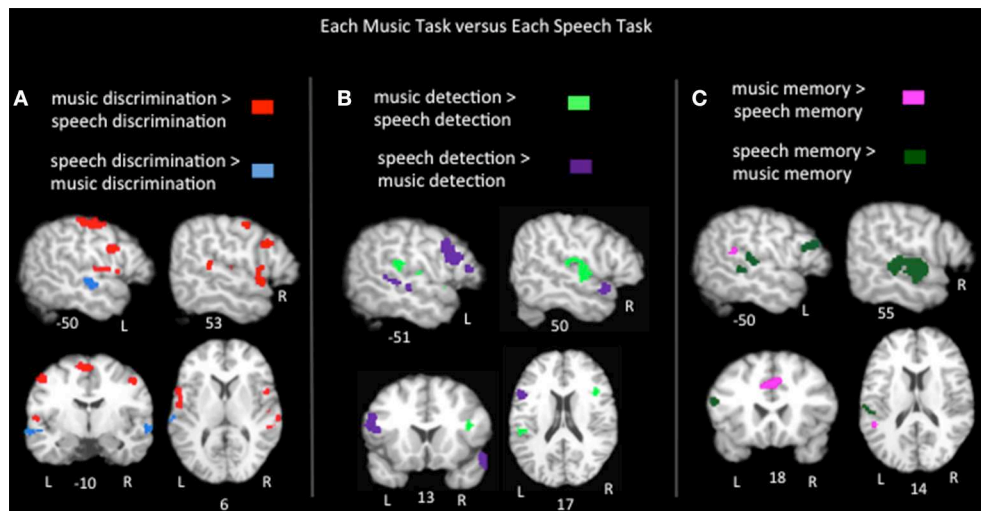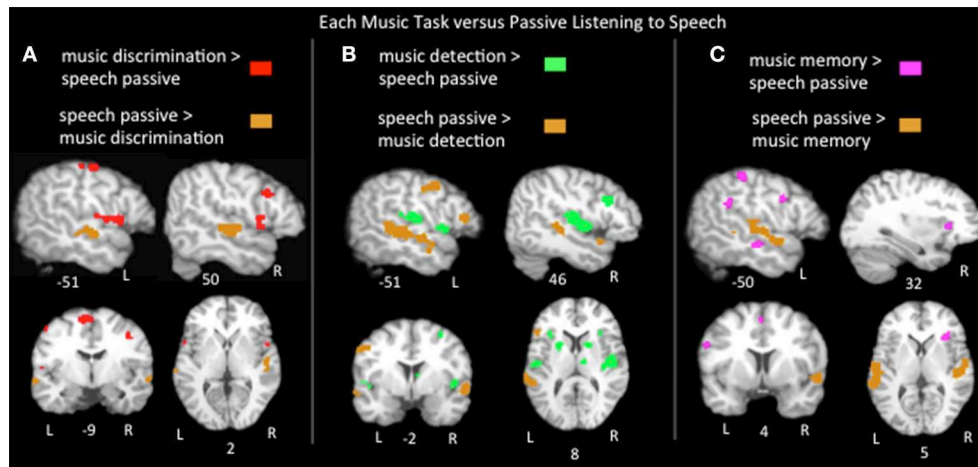

FIGURE 3 | Representative slices of the contrast results for the comparison of (A) music discrimination, (B) music error detection, and (C) music memory task conditions, compared to the corresponding speech task, $p < 0.05$, corrected.

2005; Koelsch et al., 2009; Buchsbaum et al., 2011) and are considered to be part of an auditory sensory-motor integration network (Hickok et al., 2003; Hickok and Poeppel, 2004, 2007).

Music Tasks vs. Passive Listening To Speech

Findings from various music paradigms and tasks are often reported as engaging language networks because of location; a music paradigm activating Broca's area or superior temporal regions is frequently described as recruiting classic language areas. However, it is not clear if these music paradigms are in fact engaging the language networks engaged in the natural, everyday process of listening to speech. Thus, pairwise contrasts of the ALEs for listening to speech vs. the music tasks were calculated (**Figure 4**; **Table 2**). Music discrimination > speech passive listening identified regions in bilateral precentral gyri, bilateral medial frontal gyri, left postcentral gyrus, left inferior parietal

FIGURE 4 | Representative slices of the contrast results for the comparison of (A) music discrimination, (B) music error detection, (C) music memory task conditions, compared to passive listening to speech, *p* < 0.05, corrected.

lobule, left cerebellum, right inferior and middle frontal gyri, and right superior temporal gyrus. Music error detection > speech identified bilateral precentral gyri, bilateral superior temporal gyri, bilateral insula, bilateral basal ganglia, left postcentral gyrus, left cerebellum, bilateral inferior parietal lobe, right middle frontal gyrus, right inferior frontal gyrus and the right thalamus. Music memory > speech identified portions of bilateral inferior frontal gyri, bilateral medial frontal gyri, left inferior parietal lobe, left pre and postcentral gyri, and right insula. Compared to all three music tasks, speech significantly activated bilateral superior temporal sulcus regions and only activated Broca's area (specifically the pars triangularis) more than music detection. The recruitment of Broca's area and adjacent regions for music was task dependent: compared to listening to speech, music detection and discrimination activated additional bilateral inferior precentral gyrus regions immediately adjacent to Broca's area and music memory activated the left inferior frontal gyrus more than speech (in all three subregions: pars opercularis, pars triangularis, and pars orbitalis). In the right hemisphere homolog of Broca's area, all three music tasks activated this region more than listening to speech as well as adjacent regions in the right middle frontal gyrus. All together these results suggest that the recruitment of neural resources used in speech for music processing depends on the experimental paradigm. The finding of music memory tasks eliciting widespread activation in Broca's area compared to listening to speech is likely due to the inferior frontal gyrus, and the pars opercularis in particular being consistently implicated in articulatory rehearsal and working memory (Hickok et al., 2003; Buchsbaum et al., 2011, 2005), resources that are likely recruited by the music memory tasks.
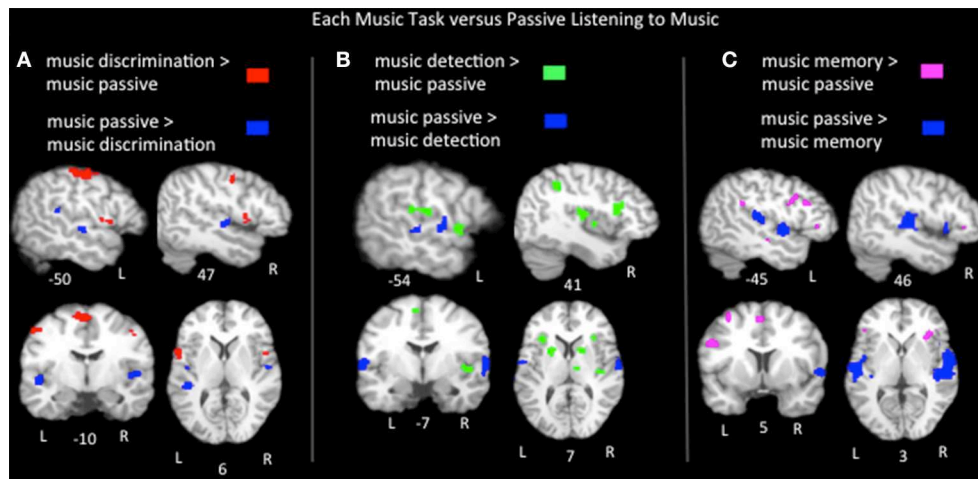
Music Tasks vs. Passive Listening To Music

Lastly we compared the music task ALEs to the music passive listening ALE using pairwise contrasts to better characterize task-specific activations to music. Results (*p* < 0.05, FDR corrected)

include: (1) music discrimination > music listening identified bilateral inferior precentral gyri, bilateral medial frontal regions, left postcentral gyrus, left inferior parietal lobule, left cerebellum, right middle frontal gyrus and right insula (2) music error detection > music listening identified bilateral medial frontal, bilateral insula, bilateral inferior parietal areas, bilateral superior temporal gyri, bilateral basal ganglia, left pre and post central gyri, right inferior and middle frontal gyri and right cerebellum; (3) music memory > passive listening identified bilateral inferior frontal gyri (pars opercularis, triangularis and orbitalis in the left hemisphere, only the latter two in the right hemisphere), bilateral medial frontal gyri, bilateral insula, bilateral cerebellum, left middle frontal gyrus, left inferior parietal lobe, left superior and middle temporal gyri, right basal ganglia, right hippocampus and right parahippocampal gyrus (**Figure 5**, **Table 2**). The medial frontal and inferior parietal activations identified in the tasks compared to listening likely reflect increased vigilance and attention due to the presence of a task, as activation in these regions is known to increase as a function of effort and performance on tasks across a variety of stimuli types and domains (Petersen and Posner, 2012; Vaden et al., 2013). To summarize the findings in Broca's area and its right hemisphere homolog, music memory tasks activated Broca's area more than just listening to music, while music discrimination and detection tasks activated right inferior frontal gyrus regions more than listening to music. Also note that all three music tasks compared to listening to music implicate regions on the anterior bank of the inferior portion of the precentral gyrus immediately adjacent to Broca's area. Significant clusters more active for music passive listening than for each of the three task conditions are found in the bilateral superior temporal gyri (**Table 2**).

Discussion

The present meta-analysis examined data from 80 functional neuroimaging studies of music and 91 studies of speech

FIGURE 5 | Representative slices of the contrast results for the comparison of (A) music discrimination, (B) music error detection, (C) music memory task conditions, compared to passive listening to music, $p < 0.05$, corrected.

to characterize the relationship between the brain networks activated by listening to speech vs. listening to music. We also compared the brain regions implicated in three frequently used music paradigms (error detection, discrimination, and memory) to the regions implicated in similar speech paradigms to determine how task effects may change how the neurobiology of music processing is related to that of speech. We replicated across a large collection of studies' previous within-subject findings that speech activates a predominately lateral temporal network, while music preferentially activates a more dorsal medial temporal network extending into the inferior parietal lobe. In Broca's area, we found overlapping resources for passive listening to speech and music in the pars opercularis, but speech "specific" resources in pars triangularis; the right hemisphere homolog of Broca's area was equally responsive to listening to speech and music. The use of a paradigm containing an explicit task (error detection, discrimination or memory) altered the relationship between the brain networks engaged in music and speech. For example, speech discrimination tasks do not activate the pars triangularis (i.e., the region identified as "speech specific" by the passive listening contrast) more than music discrimination tasks, and both speech detection and memory tasks activate the pars opercularis (i.e., the region responding equally to music and speech passive listening) more than the corresponding music tasks, while music discrimination activates pars opercularis more than speech discrimination. These findings suggest that inferior frontal contributions to music processing, and their overlap with speech resources, may be modulated by task. The following sections discuss these findings in relation to neuroanatomical models of speech and music.

Hemispheric Differences for Speech and Music

The lateralization of speech and music processing has been investigated for decades. While functional neuroimaging studies report bilateral activation for both speech and music (Jäncke et al., 2002; Abrams et al., 2011; Fedorenko et al., 2011; Rogalsky

et al., 2011), evidence from amusia, aphasia and other patient populations have traditionally identified the right hemisphere as critical for music and the left for basic language processes in most individuals (Gazzaniga, 1983; Peretz et al., 2003; Damasio et al., 2004; Hyde et al., 2006). Further evidence for hemispheric differences comes from asymmetries in early auditory cortex: left hemisphere auditory cortex has better temporal resolution and is more sensitive to rapid temporal changes critical for speech processing, while the right hemisphere auditory cortex has higher spectral resolution and is more modulated by spectral changes, which optimize musical processing (Zatorre et al., 2002; Poeppel, 2003; Schönwiesner et al., 2005; Hyde et al., 2008). Thus, left auditory cortex has been found to be more responsive to phonemes than chords, while right auditory cortex is more responsive to chords than phonemes (Tervaniemi et al., 1999, 2000). This hemispheric specialization coincides with evidence from both auditory and visual domains, suggesting that the left hemisphere tends to be tuned to local features, while the right hemisphere is tuned to more global features (Sergent, 1982; Ivry and Robertson, 1998; Sanders and Poeppel, 2007).

Hemispheric differences in the present study for speech and music vary by location. We did not find any qualitative hemispheric differences between speech and music in the temporal lobe. Speech bilaterally activated lateral superior and middle temporal regions, while music bilaterally activated more dorsal medial superior temporal regions extending into the inferior parietal lobe. However, these bilateral findings should not be interpreted as evidence against hemispheric asymmetries for speech vs. music. The hemispheric differences widely reported in auditory cortex almost always are a matter of degree, e.g., phonemes and tones both activate bilateral superior temporal regions, but a direct comparison indicates a left hemisphere preference for the speech and a right hemisphere preference for the tones (Jäncke et al., 2002; Zatorre et al., 2002). These differences would not be reflected in our ALE results because both conditions reliably activate the same regions although to

different degrees and the ALE method does not assign weight to coordinates (i.e., all the significant coordinates reported for contrasts of interest in the studies used) based on their beta or statistical values.

The frontal lobe results, however, did include some laterality differences of interest: passive listening to speech activated portions of the left inferior frontal gyrus (i.e., Broca's area), namely in the pars triangularis, significantly more than listening to music. A right inferior frontal gyrus cluster, extending into the insula, was activated significantly more for listening to music than speech. These findings in Broca's area coincide with Koelsch's neurocognitive model of music perception, in that right frontal regions are more responsive to musical stimuli and that the pars opercularis, but not the pars triangularis, is engaged in structure building of auditory stimuli (Koelsch, 2011). It is also noteworthy that the inclusion of a task altered hemispheric differences in the frontal lobes: the music discrimination tasks activated the left pars opercularis more than speech discrimination, while speech detection and memory tasks activated all of Broca's area (pars opercularis and pars triangularis) more than music detection and memory tasks; music detection and discrimination tasks, but not music memory tasks, activated the right inferior frontal gyrus more than corresponding speech tasks. These task-modulated asymmetries in Broca's area for music are particularly important when interpreting the rich electrophysiological literature of speech and music interactions. For example, both the early right anterior negativity (ERAN) and early left anterior negativity (ELAN) are modulated by speech and music, and are believed to have sources in both Broca's area and its right hemisphere homolog (Friederici et al., 2000; Maess et al., 2001; Koelsch and Friederici, 2003). Thus, the lateralization patterns found in the present study emphasize the need to consider that similar ERP effects for speech and music may arise from different underlying lateralization patterns that may be task-dependent.

Speech vs. Music in the Anterior Temporal Lobe

Superior and middle posterior temporal regions on the banks of the superior temporal sulcus were preferentially activated in each speech condition compared to each corresponding music condition in the present meta-analysis. This is not surprising, as these posterior STS regions are widely implicated in lexical semantic processing (Price, 2010) and STS regions have been found to be more responsive to syllables than tones (Jäncke et al., 2002). Perhaps more interestingly, the bilateral anterior temporal lobe (ATL) also was activated more for each speech condition than by each corresponding music condition. The role of the ATL in speech processing is debated (e.g., Scott et al., 2000 cf. Hickok and Poeppel, 2004, 2007), but the ATL is reliably sensitive to syntactic structure in speech compared to several control conditions including word lists, scrambled sentences, spectrally rotated speech, environmental sounds sequences, and melodies (Mazoyer et al., 1993; Humphries et al., 2001, 2005, 2006; Xu et al., 2005; Spitsyna et al., 2006; Rogalsky and Hickok, 2009; Friederici et al., 2010; Rogalsky et al., 2011). One hypothesis is that the ATL is implicated in combinatorial semantic processing (Wong and Gallate, 2012; Wilson et al.,

2014), although pseudoword sentences (i.e., sentences lacking meaningful content words) also activate the ATL (Humphries et al., 2006; Rogalsky et al., 2011). Several of the speech activation coordinates included in the present meta-analysis were from studies that used sentences and phrases as stimuli (with and without semantic content). It is likely that these coordinates are driving the ATL findings. Our finding that music did not activate the ATL supports the idea that the ATL is not responsive to hierarchical structure *per se* but rather needs linguistic and/or semantic information for it to be recruited.

Speech vs. Music in Broca's Area

There is no consensus regarding the role of Broca's area in receptive speech processes (e.g., Fedorenko and Kanwisher, 2011; Hickok and Rogalsky, 2011; Rogalsky and Hickok, 2011). Results from the present meta-analysis indicate that listening to speech activated both the pars opercularis and pars triangularis portions of Broca's area, while listening to music only activated the pars opercularis. The pars triangularis has been proposed to be involved in semantic integration (Hagoort, 2005) as well as in cognitive control processes such as conflict resolution (Novick et al., 2005; Rogalsky and Hickok, 2011). It is likely that the speech stimuli contain more semantic content than the music stimuli, and thus semantic integration processes may account for the speech-only response in pars triangularis. However, there was no significant difference in activations in the pars triangularis for the music discrimination and music detection tasks vs. passive listening to speech, and the music memory tasks activated portions of the pars triangularis more than listening to speech. These music task-related activations in the pars triangularis may reflect the use of semantic resources for categorization or verbalization strategies to complete the music tasks, but may also reflect increased cognitive control processes to support reanalysis of the stimuli to complete the tasks. The activation of the left pars opercularis for both speech and music replicates numerous individual studies implicating the pars opercularis in both speech and musical syntactic processing (e.g., Koelsch and Siebel, 2005; Rogalsky and Hickok, 2011) as well as in a variety of auditory working memory paradigms (e.g., Koelsch and Siebel, 2005; Buchsbaum et al., 2011).

Implications for Neuroanatomical Models of Speech and Music

It is particularly important to consider task-related effects when evaluating neuroanatomical models of the interactions between speech and music. It has been proposed that inferior frontal cortex (including Broca's area) is the substrate for shared speech-music executive function resources, such as working memory and/or cognitive control (Patel, 2003; Slevc, 2012; Slevc and Okada, 2015) as well as auditory processes such as structure analysis, repair, working memory and motor encoding (Koelsch and Siebel, 2005; Koelsch, 2011). Of particular importance here is Slevc and Okada's (2015) proposal that cognitive control may be one of the shared cognitive resources for linguistic and musical processing when reanalysis and conflict resolution is necessary. Different tasks likely recruit cognitive control resources to different degrees, and thus may explain task-related differences

for the frontal lobe's response to speech and music. There is ample evidence to support Slevc and Okada's hypothesis: classic cognitive control paradigms such as the Stroop task (Stroop, 1935; MacLeod, 1991) elicit overlapping activations in Broca's area when processing noncanonical sentence structures (January et al., 2009). Unexpected harmonic and melodic information in music interfere with Stroop task performance (Masataka and Perlovsky, 2013). The neural responses to syntactic and sentence-level semantic ambiguities in language also interact with responses to unexpected harmonics in music (Koelsch et al., 2005; Steinbeis and Koelsch, 2008b; Slevc et al., 2009; Perruchet and Poulin-Charronnat, 2013). The present results suggest that this interaction between language and music possibly via cognitive control mechanisms, localized to Broca's area, may be task driven and not inherent to the stimuli themselves. In addition, many language/music interaction studies use a reading language task with simultaneous auditory music stimuli; it is possible that a word-by-word presentation reading paradigm engages additional reanalysis mechanisms that may dissociate from resources used in auditory speech processing (Tillmann, 2012).

Slevc and Okada suggest that future studies should use tasks designed to drive activation of specific processes, presumably including reanalysis. However, the present findings suggest it is possible that these task-induced environments may actually drive overlap of neural resources for speech and music not because they are taxing shared sensory computations but rather because they are introducing additional processes that are not elicited during typical, naturalistic music listening. For example, consider the present findings in the left pars triangularis: this region is not activated during listening to music, but is activated while listening to speech. However, by presumably increasing the need for reanalysis mechanisms via discrimination or memory tasks, music does recruit this region.

There may be inferior frontal shared mechanisms that are stimulus driven while others are task driven: Broca's area is a diverse region in terms of its cytoarchitecture, connectivity and response properties (Amunts et al., 1999; Anwander et al., 2007; Rogalsky and Hickok, 2011; Rogalsky et al., in press). It is possible that some networks are task driven and some are stimulus driven. The hypotheses of Koelsch et al. are largely grounded in behavioral and electrophysiology studies that indicate an interaction between melodic and syntactic information (e.g., Koelsch et al., 2005; Fedorenko et al., 2009; Hoch et al., 2011). It is not known if these interactions are stimulus driven; a variety of tasks have been used in this literature, including discrimination, anomaly/error detection, (Koelsch et al., 2005; Carrus et al., 2013), grammatical acceptability (Patel et al., 1998a; Patel, 2008), final-word lexical decision (Hoch et al., 2011), and memory/comprehension tasks (Fedorenko et al., 2009, 2011). In addition, there is substantial variability across individual subjects, both functionally and anatomically, within Broca's area (Amunts et al., 1999; Schönwiesner et al., 2007; Rogalsky et al., in press). Thus, future within-subject studies are needed to better understand the role of cognitive control and other domain-general resources in musical processing independent of task.

Different tasks, regardless of the nature of the stimuli, may require different attentional resources (Shallice, 2003). Thus, it is possible that the inferior frontal differences between the music tasks and passive listening to music and speech are due to basic attentional differences, not the particular task *per se*. However, we find classic domain-general attention systems in the anterior cingulate and medial frontal cortex to be significantly activated across all conditions: music tasks, speech tasks, passive listening to music and passive listening to speech. These findings support Slevc and Okada's (2015) claim that domain-general attention mechanisms facilitated by anterior cingulate and medial frontal cortex are consistently engaged for music as they are for speech. Each of our music task conditions do activate these regions significantly more than the passive listening, suggesting that the midline domain-general attention mechanisms engaged by music can be further activated by explicit tasks.

Limitations and Future Directions

One issue in interpreting our results may be the proximity of distinct networks for speech and music (Peretz, 2006; Koelsch, 2011). Overlap in fMRI findings, particularly in a meta-analysis, does not necessarily mean that speech and music share resources in those locations. It is certainly possible that the spatial resolution of fMRI is not sufficient to visualize separation occurring at a smaller scale (Peretz and Zatorre, 2005; Patel, 2012). However, our findings of spatially distinct regions for music and speech clearly suggest the recruitment of distinct brain networks for speech and music.

Another potential issue related to the limitations of fMRI is that of sensitivity. Continuous fMRI scanning protocols (i.e., stimuli are presented simultaneously with the noise of scanning) and sparse temporal sampling fMRI protocols (i.e., stimuli are presented during silent periods between volume acquisitions) are both included in the present meta-analyses. It has been suggested that the loud scanner noise may reduce sensitivity to detecting hemodynamic response to stimuli, particularly complex auditory stimuli such as speech and music (Peelle et al., 2010; Elmer et al., 2012). Thus, it is possible that effects only detected by a sparse or continuous paradigm are not represented in our ALE results. However, a comparison of continuous vs. sparse fMRI sequences found no significant differences in speech activations in the frontal lobe between the pulse sequences (Peelle et al., 2010).

Priming paradigms measuring neurophysiological responses (ERP, fMRI, etc.) are one way to possibly circumvent task-related confounds in understanding the neurobiology of music in relation to that of speech. Tillmann (2012) suggests that priming paradigms may provide more insight into an individual's implicit musical knowledge than is demonstrated by performance on an explicit, overt task (e.g., Schellenberg et al., 2005; Tillmann et al., 2007). In fact, there are ERP studies that indicate that musical chords can prime processing of target words if the prime and target are semantically (i.e., emotionally) similar (Koelsch et al., 2004; Steinbeis and Koelsch, 2008a). However, most ERP priming studies investigating music or music/speech interactions have included an explicit task (e.g., Schellenberg et al., 2005; Tillmann et al., 2007; Steinbeis and Koelsch, 2008a). It is not known how

the presence of an explicit task may affect priming mechanisms via top-down mechanisms. Priming is not explored in the present meta-analysis; to our knowledge there is only one fMRI priming study of music and speech, which focused on semantic (i.e., emotion) relatedness (Steinbeis and Koelsch, 2008a).

The present meta-analysis examines networks primarily in the cerebrum. Even though almost all of the studies included in our analyses focused on cortical structures, we still identified some subcortical task-related activations: music detection compared to music passive listening activated the basal ganglia and music memory tasks activated the thalamus, hippocampus and basal ganglia compared to music passive listening. No significant differences between passive listening to speech and music were found in subcortical structures. These findings (and null results) in subcortical regions should be interpreted cautiously: given the relatively small size of these structures, activations in these areas are particularly vulnerable to spatial smoothing filters and group averaging (Raichle et al., 1991; White et al., 2001). There is also strong evidence that music and speech share subcortical resources in the brainstem (Patel, 2011), which are not addressed by the present study. For example, periodicity is a critical aspect of both speech and music and known to modulate networks between the cochlea and inferior colliculus of the brainstem (Cariani and Delgutte, 1996; Patel, 2011). Further research is needed to better understand where speech and music processing networks diverge downstream from these shared early components.

Conclusion

Listening to music and listening to speech engage distinct temporo-parietal cortical networks but share some inferior and medial frontal resources (at least at the resolution of fMRI). However, the recruitment of inferior frontal speech-processing regions for music is modulated by task. The present findings highlight the need to consider how task effects may be interacting with conclusions regarding the neurobiology of speech and music.

Acknowledgments

References

Abrams, D. A., Bhatara, A., Ryali, S., Balaban, E., Levitin, D. J., and Menon, V. (2011). Decoding temporal structure in music and speech relies on shared brain resources but elicits different fine-scale spatial patterns. *Cereb. Cortex* 21, 1507–1518. doi: 10.1093/cercor/bhq198

Adank, P. (2012). Design choices in imaging speech comprehension: an activation likelihood estimation (ALE) meta-analysis. *Neuroimage* 63, 1601–1613. doi: 10.1016/j.neuroimage.2012.07.027

Amunts, K., Schleicher, A., Bürgel, U., Mohlberg, H., Uylings, H. B. M., and Zilles, K. (1999). Broca's region revisited: cytoarchitecture and intersubject variability. *J. Comp. Neurol.* 412, 319–341.

Anwander, A., Tittgemeyer, M., von Cramon, D. Y., Friederici, A. D., and Knösche, T. R. (2007). Connectivity-based parcellation of Broca's area. *Cereb. Cortex* 17, 816–825. doi: 10.1093/cercor/bhk034

Baker, E., Blumstein, S. E., and Goodglass, H. (1981). Interaction between phonological and semantic factors in auditory comprehension. *Neuropsychology* 19, 1–15. doi: 10.1016/0028-3932(81)90039-7

Basso, A., and Capitani, E. (1985). Spared musical abilities in a conductor with global aphasia and ideomotor apraxia. *J. Neurol. Neurosurg. Psychiatry* 48, 407–412. doi: 10.1136/jnnp.48.5.407

Besson, M., Chobert, J., and Marie, C. (2011). Transfer of training between music and speech: common processing, attention, and memory. *F. Psychol.* 2:94. doi: 10.3389/fpsyg.2011.00094

Besson, M., and Faita, F. (1995). An event-related potential (ERP) study of musical expectancy: comparison of musicians with nonmusicians. *J. Exp. Psychol. Hum. Percept. Perform.* 21, 1278–1296. doi: 10.1037/0096-1523.21.6.1278

Besson, M., and Schön, D. (2001). Comparison between language and music. *Ann. N.Y. Acad. Sci.* 930, 232–258. doi: 10.1111/j.1749-6632.2001.tb05736.x

Brattico, E., Tervaniemi, M., Näätänen, R., and Peretz, I. (2006). Musical scale properties are automatically processed in the human auditory cortex. *Brain Res.* 1117, 162–174. doi: 10.1016/j.brainres.2006.08.023

Buchsbaum, B. R., Baldo, J., Okada, K., Berman, K. F., Dronkers, N., D'Esposito, M., et al. (2011). Conduction aphasia, sensory-motor integration, and phonological short-term memory - an aggregate analysis of lesion and fMRI data. *Brain Lang.* 119, 119–128. doi: 10.1016/j.bandl.2010.12.001

Buchsbaum, B. R., Olsen, R. K., Koch, P., and Berman, K. F. (2005). Human dorsal and ventral auditory streams subserve rehearsal-based and echoic processes during verbal working memory. *Neuron* 48, 687–697. doi: 10.1016/j.neuron.2005.09.029

Cant, J. S., and Goodale, M. A. (2007). Attention to form or surface properties modulates different regions of human occipitotemporal cortex. *Cereb. Cortex* 17, 713–731. doi: 10.1093/cercor/bhk022

Cariani, P. A., and Delgutte, B. (1996). Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. *J. Neurophysiol.* 76, 1698–1716.

Carrus, E., Pearce, M. T., and Bhattacharya, J. (2013). Melodic pitch expectation interacts with neural responses to syntactic but not semantic violations. *Cortex* 49, 2186–2200. doi: 10.1016/j.cortex.2012.08.024

Chawla, D., Rees, G., and Friston, K. J. (1999). The physiological basis of attentional modulation in extrastriate visual areas. *Nat. Neurosci.* 2, 671–676. doi: 10.1038/10230

Corbetta, M., Miezin, F. M., Dobmeyer, S., Shulman, G. L., and Petersen, S. E. (1990). Attentional modulation of neural processing of shape, color, and velocity in humans. *Science* 248, 1556–1559. doi: 10.1126/science.2360050

Damasio, H., Tranel, D., Grabowski, T., Adolphs, R., and Damasio, A. (2004). Neural systems behind word and concept retrieval. *Cognition* 92, 179–229. doi: 10.1016/j.cognition.2002.07.001

Dronkers, N. F., Wilkins, D. P., Van Valin, R. D. Jr., Redfern, B. B., and Jaeger, J. J. (2004). Lesion analysis of the brain areas involved in language comprehension. *Cognition* 92, 145–177. doi: 10.1016/j.cognition.2003.11.002

Eickhoff, S. B., Bzdok, D., Laird, A. R., Kurth, F., and Fox, P. T. (2012). Activation likelihood estimation revisited. *Neuroimage* 59, 2349–2361. doi: 10.1016/j.neuroimage.2011.09.017

Eickhoff, S. B., Laird, A. R., Grefkes, C., Wang, L. E., Zilles, K., and 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. *Hum. Brain Mapp.* 30, 2907–2926. doi: 10.1002/hbm.20718

Elmer, S., Meyer, S., and Jäncke, L. (2012). Neurofunctional and behavioral correlates of phonetic and temporal categorization in musically trained and untrained subjects. *Cereb. Cortex* 22, 650–658. doi: 10.1093/cercor/bhr142

Fedorenko, E., Behr, M. K., and Kanwisher, N. (2011). Functional specificity for high-level linguistic processing in the human brain. *Proc. Natl. Acad. Sci. U.S.A.* 108, 16428–16433. doi: 10.1073/pnas.1112937108

Fedorenko, E., and Kanwisher, N. (2011). Some regions within Broca's area do respond more strongly to sentences than to linguistically degraded stimuli: a comment on Rogalsky and Hickok (2010). *J. Cogn. Neurosci.* 23, 2632–2635. doi: 10.1162/jocn_a_00043

Fedorenko, E., Patel, A., Casasanto, D., Winawer, J., and Gibson, E. (2009). Structural integration in language and music: evidence for a shared system. *Mem. Cogn.* 37, 1–9. doi: 10.3758/MC.37.1.1

Frances, R., Lhermitte, F., and Verdy, M. F. (1973). Le deficit musical des aphasiques. *Appl. Psychol.* 22, 117–135. doi: 10.1111/j.1464-0597.1973.tb00391.x

Friederici, A. D., Kotz, S. A., Scott, S. K., and Obleser, J. (2010). Disentangling syntax and intelligibility in auditory language comprehension. *Hum. Brain Mapp.* 31, 448–457. doi: 10.1002/hbm.20878

Friederici, A. D., Wang, Y., Herrmann, C. S., Maess, B., and Oertel, U. (2000). Localization of early syntactic processes in frontal and temporal cortical areas: a magnetoencephalographic study. *Hum. Brain Mapp.* 11, 1–11. doi: 10.1002/1097-0193(200009)11:1<1::AID-HBM10>3.0.CO;2-B

Gazzaniga, M. S. (1983). Right hemisphere language following brain bisection: a 20-year perspective. *Am. Psychol.* 38, 525–537. doi: 10.1037/0003-066X.38.5.525

Geiser, E., Zaehle, T., Jancke, L., and Meyer, M. (2008). The neural correlate of speech rhythm as evidenced by metrical speech processing. *J. Cogn. Neurosci.* 20, 541–552. doi: 10.1162/jocn.2008.20029

Grahn, J. A. (2012). "Advances in neuroimaging techniques: Implications for the shared syntactic integration resource hypothesis," in *Language and Music as Cognitive Systems*, eds P. Rebuschat, M. Rohrmeier, J. Hawkins, and I. Cross (Oxford: Oxford University Press), 235–241.

Hagoort, P. (2005). On Broca, brain and binding: a new framework. *Trends Cogn. Sci.* 9, 416–423. doi: 10.1016/j.tics.2005.07.004

Henschen, S. E. (1924). On the function of the right hemisphere of the brain in relation to the left in speech, music and calculation. *Brain* 44, 110–123.

Hickok, G., Buchsbaum, B., Humphries, C., and Muftuler, T. (2003). Auditory-motor interaction revealed by fMRI: speech, music, and working memory in area Spt. *J. Cogn. Neurosci.* 15, 673–682. doi: 10.1162/089892903322307393

Hickok, G., and Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition* 92, 67–99. doi: 10.1016/j.cognition.2003.10.011

Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402. doi: 10.1038/nrn2113

Hickok, G., and Rogalsky, C. (2011). What does Broca's area activation to sentences reflect? *J. Cogn. Neurosci.* 23, 2629–2631. doi: 10.1162/jocn_a_00044

Hoch, L., Poulin-Charronnat, B., and Tillmann, B. (2011). The influence of task-irrelevant music on language processing: syntactic and semantic structures. *Front. Psychol.* 2:112. doi: 10.3389/fpsyg.2011.00112

Humphries, C., Binder, J. R., Medler, D. A., and Liebenthal, E. (2006). Syntactic and semantic modulation of neural activity during auditory sentence comprehension. *J. Cogn. Neurosci.* 18, 665–679. doi: 10.1162/jocn.2006.18.4.665

Humphries, C., Love, T., Swinney, D., and Hickok, G. (2005). Response of anterior temporal cortex to syntactic and prosodic manipulations during sentence processing. *Hum. Brain Mapp.* 26, 128–138. doi: 10.1002/hbm.20148

Humphries, C., Willard, K., Buchsbaum, B., and Hickok, G. (2001). Role of anterior temporal cortex in auditory sentence comprehension: an fMRI study. *Neuroreport* 12, 1749–1752. doi: 10.1097/00001756-200106130-00046

Hyde, K. L., Peretz, I., and Zatorre, R. J. (2008). Evidence for the role of the right auditory cortex in fine pitch resolution. *Neuropsychologia* 46, 632–639. doi: 10.1016/j.neuropsychologia.2007.09.004

Hyde, K. L., Zatorre, R. J., Griffiths, T. D., Lerch, J. P., and Peretz, I. (2006). Morphometry of the amusic brain: a two-site study. *Brain* 129, 2562–2570. doi: 10.1093/brain/awl204

Ivry, R. B., and Robertson, L. C. (1998). *The Two Sides of Perception*. Cambridge, MA: MIT Press.

Jäncke, L. (2008). Music, memory and emotion. *J. Biol.* 7, 21. doi: 10.1186/jbiol82

Jäncke, L., Wüstenberg, T., Scheich, H., and Heinze, H. J. (2002). Phonetic perception and the temporal cortex. *Neuroimage* 15, 733–746. doi: 10.1006/nimg.2001.1027

January, D., Trueswell, J. C., and Thompson-Schill, S. L. (2009). Co-localization of stroop and syntactic ambiguity resolution in Broca's area: implications for the neural basis of sentence processing. *J. Cogn. Neurosci.* 21, 2434–2444. doi: 10.1162/jocn.2008.21179

Koelsch, S. (2005). Neural substrates of processing syntax and semantics in music. *Curr. Opin. Neurobiol.* 15, 207–212. doi: 10.1016/j.conb.2005.03.005

Koelsch, S. (2011). Toward a neural basis of music perception – a review and updated model. *Front. Psychol.* 2:110. doi: 10.3389/fpsyg.2011.00110

Koelsch, S., and Friederici, A. D. (2003). Toward the neural basis of processing structure in music. *Ann. N.Y. Acad. Sci.* 999, 15–28. doi: 10.1196/annals.1284.002

Koelsch, S., Gunter, T. C., Wittfoth, M., and Sammler, D. (2005). Interaction between syntax processing in language and in music: an ERP study. *J. Cogn. Neurosci.* 17, 1565–1577. doi: 10.1162/089892905774597290

Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., and Friederici, A. D. (2004). Music, language and meaning: brain signatures of semantic processing. *Nat. Neurosci.* 7, 302–307. doi: 10.1038/nn1197

Koelsch, S., and Siebel, W. A. (2005). Towards a neural basis of music perception. *Trends Cogn. Sci.* 9, 578–584. doi: 10.1016/j.tics.2005.10.001

Koelsch, S., Schulze, K., Sammler, D., Fritz, T., Müller, K., and Gruber, O. (2009). Functional architecture of verbal and tonal working memory: an fMRI study. *Hum. Brain Mapp.* 30, 859–873. doi: 10.1002/hbm.20550

Luria, A. R., Tsvetkova, L., and Futer, D. S. (1965). Aphasia in a composer. *J. Neurol. Sci.* 1, 288–292. doi: 10.1016/0022-510X(65)90113-9

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychol. Bull.* 109, 163–203. doi: 10.1037/0033-2909.109.2.163

Maess, B., Koelsch, S., Gunter, T. C., and Friederici, A. D. (2001). Musical syntax is processed in Broca's area: an MEG study. *Nat. Neurosci.* 4, 540–545. doi: 10.1038/87502

Maillard, L., Barbeau, E. J., Baumann, C., Koessler, L., Bénar, C., Chauvel, P., et al. (2011). From perception to recognition memory: time course and lateralization of neural substrates of word and abstract picture processing. *J. Cogn. Neurosci.* 23, 782–800. doi: 10.1162/jocn.2010.21434

Masataka, N., and Perlovsky, L. (2013). Cognitive interference can be mitigated by consonant music and facilitated by dissonant music. *Sci. Rep.* 3, 1–6. doi: 10.1038/srep02028

Mazoyer, B. M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., et al. (1993). The cortical representation of speech. *J. Cogn. Neurosci.* 5, 467–479. doi: 10.1162/jocn.1993.5.4.467

Ni, W., Constable, R. T., Mencl, W. E., Pugh, K. R., Fulbright, R. K., Shaywitz, S. E., et al. (2000). An event-related neuroimaging study distinguishing form and content in sentence processing. *J. Cogn. Neurosci.* 12, 120–133. doi: 10.1162/08989290051137648

Noesselt, T., Shah, N. J., and Jäncke, L. (2003). Top-down and bottom-up modulation of language related areas- an fMRI study. *BMC Neurosci.* 4:13. doi: 10.1186/1471-2202-4-13

Novick, J. M., Trueswell, J. C., and Thompson-Schill, S. L. (2005). Cognitive control and parsing: reexamining the role of Broca's area in sentence comprehension. *Cogn. Affect. Behav. Neurosci.* 5, 263–281. doi: 10.3758/CABN.5.3.263

Oechslin, M. S., Meyer, M., and Jäncke, L. (2010). Absolute pitch – functional evidence of speech-relevant auditory acuity. *Cereb. Cortex* 20, 447–455. doi: 10.1093/cercor/bhp113

Patel, A. (2003). Language, music, syntax and the brain. *Nat. Neurosci.* 6, 674–681. doi: 10.1038/nn1082

Patel, A. D. (2005). The relationship of music to the melody of speech and to syntactic processing disorders in aphasia. *Ann. N.Y. Acad. Sci.* 1060, 59–70. doi: 10.1196/annals.1360.005

Patel, A. D. (2008). *Music, Language, and the Brain*. New York, NY: Oxford Univ. Press.

Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Front. Psychol.* 2:142. doi: 10.3389/fpsyg.2011.00142

Patel, A. D. (2012). "Language, music, and the brain: a resource-sharing framework," in *Language and Music as Cognitive Systems*, eds P. Rebuschat, M. Rohrmeier, J. Hawkins, and I. Cross (New York, NY: Oxford University Press), 204–223.

Patel, A. D. (2013). "Sharing and nonsharing of brain resources for language and music," in *Language, Music, and the Brain*, ed M. Arbib (Cambridge, MA: MIT Press), 329–355.

Patel, A. D., Gibson, E., Ratner, J., Besson, M., and Holcomb, P. (1998a). Processing syntactic relations in language and music: an event-related potential study. *J. Cogn. Neurosci.* 10, 717–733. doi: 10.1162/0898929985 63121

Patel, A. D., Peretz, I., Tramo, M., and Labreque, R. (1998b). Processing prosodic and musical patterns: a neuropsychological investigation. *Brain Lang.* 61, 123–144. doi: 10.1006/brln.1997.1862

Peelle, J. E., Eason, R. J., Schmitter, S., Schwarzbauer, C., and Davis, M. H. (2010). Evaluating an acoustically quiet EPI sequence for use in fMRI studies of speech and auditory processing. *Neuroimage* 52, 1410–1419. doi: 10.1016/j.neuroimage.2010.05.015

Peretz, I. (2006). The nature of music from a biological perspective. *Cognition* 100, 1–32. doi: 10.1016/j.cognition.2005.11.004

Peretz, I., Belleville, S., and Fontaine, S. (1997). Dissociations between music and language functions after cerebral resection: a new case of amusia without aphasia. *Can. J. Exp. Psychol.* 51, 354–368. doi: 10.1037/1196-1961. 51.4.354

Peretz, I., Champod, A. S., and Hyde, K. (2003). Varieties of musical disorders. *Ann. N.Y. Acad. Sci.* 999, 58–75. doi: 10.1196/annals.1284.006

Peretz, I. and Hyde, K. L. (2003). What is specific to music processing? Insights from congenital amusia. *Trends. Cog. Sci.* 7, 362–367. doi: 10.1016/S1364-6613(03)00150-5

Peretz, I., Kolinsky, R., Tramo, M., Labrecque, R., Hublet, C., Demeurisse, G., et al. (1994). Functional dissociations following bilateral lesions of auditory cortex. *Brain* 117, 1283–1302. doi: 10.1093/brain/117.6.1283

Peretz, I., and Zatorre, R. J. (2005). Brain organization for music processing. *Annu. Rev. Psychol.* 56, 89–114. doi: 10.1146/annurev.psych.56.091103.070225

Perruchet, P., and Poulin-Charronnat, B. (2013). Challenging prior evi- dence for a shared syntactic processor for language and music. *Psychon. Bull. Rev.* 20, 310–317. doi: 10.3758/s13423-012-0344-5

Petersen, S. E., and Posner, M. I. (2012). The attention system of the human brain: 20 years later. *Annu. Rev. Neurosci.* 35, 73–89. doi: 10.1146/annurev-neuro-062111-150525

Platel, H., Price, C., Baron, J. C., Wise, R., Lambert, J., Frackowiak, R. S., et al. (1997). The structural components of music perception: a functional anatomical study. *Brain* 120, 229–243. doi: 10.1093/brain/120. 2.229

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as 'assymetric sampling in time. *Speech Commun.* 41, 245–255. doi: 10.1016/S0167-6393(02)00107-3

Price, C. J. (2010). The anatomy of language: a review of 100 fMRI studies published in 2009. *Ann. N.Y. Acad. Sci.* 1191, 62–88. doi: 10.1111/j.1749-6632.2010.05444.x

Raichle, M. E., Mintun, M. A., Shertz, L. D., Fusselman, M. J., and Miezen, F. (1991). The influence of anatomical variability on functional brain mapping with PET: a study of intrasubject versus intersubject averaging. *J. Cereb. Blood Flow Metab.* 11, S364.

Rogalsky, C., Almeida, D., Sprouse, J., and Hickok, G. (in press). Sentence processing selectivity in Broca's area: evident for structure but not syntactic movement. *Lang. Cogn. Neurosci.*

Rogalsky, C., and Hickok, G. (2009). Selective attention to semantic and syntactic features modulates sentence processing networks in anterior temporal cortex. *Cereb. Cortex* 19, 786–796. doi: 10.1093/cercor/bhn126

Rogalsky, C., and Hickok, G. (2011). The role of Broca's area in sentence comprehension. *J. Cogn. Neurosci.* 23, 1664–1680. doi: 10.1162/jocn.2010.21530

Rogalsky, C., Poppa, N., Chen, K. H., Anderson, S. W., Damasio, H., Love, T., et al. (2015). Speech repetition as a window on the neurobiology of auditory-motor integration for speech: a voxel-based lesion symptom mapping study. *Neuropsychologia* 71, 18–27. doi: 10.1016/j.neuropsychologia.2015. 03.012

Rogalsky, C., Rong, F., Saberi, K. and Hickok, G. (2011). Functional anatomy of language and music perception: temporal and structural factors investigated using functional magnetic resonance imaging. *J. Neurosci.* 31, 3843–3852. doi: 10.1523/JNEUROSCI/4515-10.2011

Rorden, C., and Brett, M. (2000). Stereotaxic display of brain lesions. *Behav. Neurol.* 12, 191–200. doi: 10.1155/2000/421719

Sammler, D., Koelsch, S., and Friederici, A. D. (2011). Are left fronto- temporal brain areas a prerequisite for normal music-syntactic pro- cessing? *Cortex* 47, 659–673. doi: 10.1016/j.cortex.2010.04.007

Sanders, L. D., and Poeppel, D. (2007). Local and global auditory processing: behavioral and ERP evidence. *Neuropsychologia* 45, 1172–1186. doi: 10.1016/j.neuropsychologia.2006.10.010

Scheich, H., Brechmann, A., Brosch, M., Budinger, E., and Ohl, F. W. (2007). The cognitive auditory cortex: task-specificity of stimulus representations. *Hear. Res.* 229, 213–224. doi: 10.1016/j.heares.2007.01.025

Schellenberg, E. G., Bigand, E., Poulin-Charronnat, B., Garnier, C., and Stevens, C. (2005). Children's implicit knowledge of harmony in Western music. *Dev. Sci.* 8, 551–566. doi: 10.1111/j.1467-7687.2005.00447.x

Schönwiesner, M., Novitski, N., Pakarinen, S., Carlson, S., Tervaniemi, M., and Näätänen, R. (2007). Heschl's gyrus, posterior superior temporal gryus, and mid-ventrolateral prefrontal cortex have different roles in the detection of acoustic changes. *J. Neurophysiol.* 97, 2075–2082. doi: 10.1152/jn. 01083.2006

Schönwiesner, M., Rübsamen, R., and von Cramon, D. Y. (2005). Hemispheric asymmetry for spectral and temporal processing in the human antero−lateral auditory belt cortex. *Eur. J. Neurosci.* 22, 1521–1528. doi: 10.1111/j.1460-9568.2005.04315.x

Schwartz, M. F., Faseyitan, O., Kim, J., and Coslett, H. B. (2012). The dorsal stream contribution to phonological retrieval in object naming. *Brain* 135(Pt 12), 3799–3814. doi: 10.1093/brain/aws300

Scott, S. K., Blank, C. C., Rosen, S., and Wise, R. J. S. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain* 123, 2400–2406. doi: 10.1093/brain/123.12.2400

Sergent, J. (1982). About face: left-hemisphere involvement in processing phsyiognomies. *J. Exp. Psychol. Hum. Percept. Perform.* 8, 1–14.

Shallice, T. (2003). Functional imaging and neuropsychology findings: how can they be linked? *Neuroimage* 20, S146–S154. doi: 10.1016/j.neuroimage.2003.09.023

Slevc, R. L. (2012). Language and music: sound, structure and meaning. *WIREs Cogn. Sci.* 3, 483–492. doi: 10.1002/wcs.1186

Slevc, L. R., Rosenberg, J. C., and Patel, A. D. (2009). Making psycholinguistics musical: self-paced reading time evidence for shared processing of linguistic and musical syntax. *Psychon. Bull. Rev.* 16, 374–381. doi: 10.3758/16.2.374

Slevc, L. R., and Okada, B. M. (2015). Processing structure in language and music: a case for shared reliance on cognitive control. *Psychon. Bull. Rev.* 22, 637–652. doi: 10.3758/s13423-014-0712-4

Specht, K., Willmes, K., Shah, N. J., and Jäncke, L. (2003). Assessment of reliability in functional imaging studies. *J. Magn. Reson. Imag.* 17, 463–471. doi: 10.1002/jmri.10277

Spitsyna, G., Warren, J. E., Scott, S. K., Turkheimer, F. E., and Wise, R. J. (2006). Converging language streams in the human temporal lobe. *J. Neurosci.* 26, 7328–7336. doi: 10.1523/JNEUROSCI.0559-06.2006

Steinbeis, N., and Koelsch, S. (2008a). Comparing the processing of music and language meaning using EEG and fMRI provides evidence for similar and distinct neural representations. *PLoS ONE.* 3:e2226. doi: 10.1371/journal.pone.0002226

Steinbeis, N., and Koelsch, S. (2008b). Shared neural resources between music and language indicate semantic processing of musical tension-resolution patterns. *Cereb. Cortex* 18, 1169–1178. doi: 10.1093/cercor/bhm149

Steinke, W. R., Cuddy, L. L., and Holden, R. R. (1997). Dissociation of musical tonality and pitch memory from nonmusical cognitive abilities. *Can. J. Exp. Psychol.* 51:316. doi: 10.1037/1196-1961.51.4.316

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18, 643–662. doi: 10.1037/h0054651

Tervaniemi, M. (2001). Musical sound processing in the human brain. Evidence from electric and magnetic recordings. *Ann. N.Y. Acad. Sci.* 930, 259–272. doi: 10.1111/j.1749-6632.2001.tb05737.x

Tervaniemi, M., and Hugdahl, K. (2003). Lateralization of auditory-cortex functions. *Brain Res. Brain Res. Rev.* 43, 231–246. doi: 10.1016/j.brainresrev.2003.08.004

Tervaniemi, M., Kujala, A., Alho, K., Virtanen, J., Ilmoniemi, R., and Näätänen, R. (1999). Functional specialization of the human auditory cortex in processing

phonetic and musical sounds: a magnetoencephalographic (MEG) study. *Neuroimage* 9, 330–336. doi: 10.1006/nimg.1999.0405

Tervaniemi, M., Mendvedev, S. V., Alho, K., Pakhomov, S. V., Roudas, M. S., Van Zuijen, T. L., et al. (2000). Lateralized automatic auditory processing of phonetic versus musical information: a PET study. *Hum. Brain Mapp.* 10, 74–79. doi: 10.1002/(SICI)1097-0193(200006)10:2<74::AID-HBM30>3.0.CO;2-2

Tillmann, B. (2012). Music and language perception: expectations, structural integration, and cognitive sequencing. *Top. Cogn. Sci.* 4, 568–584. doi: 10.1111/j.1756-8765.2012.01209.x

Tillmann, B., Peretz, I., Bigand, E., and Gosselin, N. (2007). Harmonic priming in an amusic patient: the power of implicit tasks. *Cogn. Neuropsychol.* 24, 603–622. doi: 10.1080/02643290701609527

Turkeltaub, P. E., and Coslett, H. B. (2010). Localization of sublexical speech perception components. *Brain Lang.* 114, 1–15. doi: 10.1016/j.bandl.2010.03.008

Turkeltaub, P. E., Eickhoff, S. B., Laird, A. R., Fox, M., Wiener, M., and Fox, P. (2012). Minimizing within−experiment and within−group effects in activation likelihood estimation meta−analyses. *Hum. Brain Mapp.* 33, 1–13. doi: 10.1002/hbm.21186

Tzortzis, C., Goldblum, M. C., Dang, M., Forette, F., and Boller, F. (2000). Absence of amusia and preserved naming of musical instruments in an aphasic composer. *Cortex* 36, 227–242. doi: 10.1016/S0010-9452(08)70526-4

Vaden, K. I. Jr., Kuchinsky, S. E., Cute, S. L., Ahlstrom, J. B., Dubno, J. R., and Eckert, M. A. (2013). The cingulo-opercular network provides word-recognition benefit. *J. Neurosci.* 33, 18979–18986. doi: 10.1523/JNEUROSCI.1417-13.2013

Von Kriegstein, K., Eiger, E., Kleinschmidt, A., and Giraud, A. L. (2003). Modulation of neural responses to speech by directing attention to voices or verbal content. *Cogn. Brain Res.* 17, 48–55. doi: 10.1016/S0926-6410(03)00079-X

White, T., O'Leary, D., Magnotta, V., Arndt, S., Flaum, M., and Andreasen, N. C. (2001). Anatomic and functional variability: the effects of filter size in group fMRI data analysis. *Neuroimage* 13, 577–588. doi: 10.1006/nimg.2000.0716

Wilson, S. M., DeMarco, A. T., Henry, M. L., Gesierich, B., Babiak, M., Mandelli, M. L., et al. (2014). What role does the anterior temporal lobe play in sentence-level processing? Neural correlates of syntactic processing in semantic variant primary progressive aphasia. *J. Cogn. Neurosci.* 26, 970–985. doi: 10.1162/jocn_a_00550

Wong, C., and Gallate, J. (2012). The function of the anterior temporal lobe: a review of the empirical evidence. *Brain Res.* 1449, 94–116. doi: 10.1016/j.brainres.2012.02.017

Xu, J., Kemeny, S., Park, G., Frattali, C., and Braun, A. (2005). Language in context: emergent features of word, sentence, and narrative comprehension. *Neuroimage* 25, 1002–1015. doi: 10.1016/j.neuroimage.2004.12.013

Yamadori, A., Osumi, Y., Masuhara, S., and Okubo, M. (1977). Preservation of singing in Broca's aphasia. *J. Neurol. Neurosurg. Psychiatry* 40, 221–224. doi: 10.1136/jnnp.40.3.221

Zatorre, R. J., Belin, P., and Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends Cogn. Sci.* 6, 37–46. doi: 10.1016/S1364-6613(00)01816-7

Zatorre, R. J., and Gandour, J. T. (2008). Neural specializations for speech and pitch: moving beyond the dichotomies. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 1087–1104. doi: 10.1098/rstb.2007.2161

Zheng, Z. Z., Munhall, K. G., and Johnsrude, I. S. (2010). Functional overlap between regions involved in speech perception and in monitoring one's own voice during speech production. *J. Cogn. Neurosci.* 22, 1770–1781. doi: 10.1162/jocn.2009.21324