



Stimulus complexity and categorical effects in human auditory cortex: an Activation Likelihood Estimation meta-analysis

Fabienne Samson^{1,2*}, Thomas A. Zeffiro³, Alain Toussaint^{1,2} and Pascal Belin^{4,5}

¹ Centre d'Excellence en Troubles Envahissants du Développement de l'Université de Montréal, Montréal, QC, Canada

² Centre de Recherche Fernand-Seguin, Department of Psychiatry, Université de Montréal, Montréal, QC, Canada

³ Neural Systems Group, Massachusetts General Hospital, Boston, MA, USA

⁴ Institute of Neuroscience and Psychology, University of Glasgow, Glasgow, Scotland, UK

⁵ International Laboratories for Brain, Music and Sound, McGill University and Université de Montréal, Montréal, QC, Canada

Edited by:

Josef P. Rauschecker, Georgetown University School of Medicine, USA

Reviewed by:

Jonas Obleser, Max Planck Institute for Human Cognitive and Brain Sciences, Germany

Josef P. Rauschecker, Georgetown University School of Medicine, USA

*Correspondence:

Fabienne Samson, Hôpital Rivière-des-Prairies, Recherche TN, 7070 Boulevard Perras, Montréal, QC, Canada H1E 1A4.
e-mail: samsonfabienne1@gmail.com

Investigations of the functional organization of human auditory cortex typically examine responses to different sound categories. An alternative approach is to characterize sounds with respect to their amount of variation in the time and frequency domains (i.e., spectral and temporal complexity). Although the vast majority of published studies examine contrasts between discrete sound categories, an alternative complexity-based taxonomy can be evaluated through meta-analysis. In a quantitative meta-analysis of 58 auditory neuroimaging studies, we examined the evidence supporting current models of functional specialization for auditory processing using grouping criteria based on either categories or spectro-temporal complexity. Consistent with current models, analyses based on typical sound categories revealed hierarchical auditory organization and left-lateralized responses to speech sounds, with high speech sensitivity in the left anterior superior temporal cortex. Classification of contrasts based on spectro-temporal complexity, on the other hand, revealed a striking within-hemisphere dissociation in which caudal-lateral temporal regions in auditory cortex showed greater sensitivity to spectral changes, while anterior superior temporal cortical areas were more sensitive to temporal variation, consistent with recent findings in animal models. The meta-analysis thus suggests that spectro-temporal acoustic complexity represents a useful alternative taxonomy to investigate the functional organization of human auditory cortex.

Keywords: fMRI, category, time, frequency, hierarchy

INTRODUCTION

Current accounts of the functional organization of auditory cortex, mostly based on response specificity to different sound categories, describe an organizational structure that is both hierarchical and hemispherically specialized (Rauschecker, 1998; Zatorre et al., 2002; Hackett, 2008; Rauschecker and Scott, 2009; Woods and Alain, 2009; Recanzone and Cohen, 2010).

Characterizing responses to stimuli from typical auditory categories such as music, voices, animal, or environmental sounds have provided important information about the cortical specialization for auditory processing. However, this classification may not fully account for the range of stimulus variability encountered across neuroimaging studies, as most stimuli do not fit neatly into one auditory category. For instance, an amplitude modulated tone can vary in ways that cannot be adequately characterized using typical categories. However, its characteristics can easily be described in terms of variations in time (temporal dimension) and frequency (spectral dimension), suggesting an alternative approach to stimulus classification. Accordingly, any auditory stimulus can be described with respect to its sound complexity characteristics specified with respect to changes in time and frequency. This approach represents a comprehensive characterization of sounds that is not limited to specific categories. Therefore, complexity might represent an alternative organizing principle along which to represent auditory cortical response specialization. In this conceptualization, a single

frequency sinusoidal wave (pure tone), constant over time, can be classified as simple, and a sound containing multiple components can be classified as complex with respect to the frequency domain. Examples of sounds with high spectral complexity are musical notes or sustained vowels. Similarly, a sound with acoustical structure varying over time can be classified as complex with respect to the time domain. Examples of stimuli with high temporal complexity are frequency or amplitude modulated sounds or sound sequences. Natural sounds can be complex with regards to both their frequency composition and temporal variation. Phonemes, the basic units of speech, contain multiple frequency components, called formants, which may be combined over time to produce syllables and words. Similarly, musical sequences are composed of complex changes in fundamental frequency and harmonic structure that unfold over time. Additionally, speech processing is mainly dependent on temporal information (Shannon et al., 1995), while spectral composition is most relevant for music perception (Warrier and Zatorre, 2002). Hence, acoustic complexity is not independent of sound categories and the two classification methods explored here should not be considered as mutually exclusive.

As previously proposed, an auditory stimulus can be categorized in more than one way; either based on *a priori* knowledge about the characterizing features of the sound source or on the basis of a sound's acoustic pattern in the frequency and time domain (Griffiths and Warren, 2004). Additionally, some studies suggest

that auditory cortex activation to sounds of a given category could reflect a specialized response to the acoustic components characterizing sounds within this category (Lewis et al., 2005, 2009). This suggests a certain level of interaction between the cortical processes involved in the analysis of acoustic features and those showing sensitivity to sound categories. However, recently Leaver and Rauschecker (2010) demonstrated categorical effects of speech and music stimuli even when controlling for changes in spectral and temporal dimensions. The two classification approaches are therefore not mutually exclusive and both methods seem relevant and can complement each other in revealing different aspects of cortical auditory specialization. In vision, cortical representation of stimulus complexity has been described with simple (first-order) information being analyzed within primary visual cortex (V1) and complex (second-order) information processing involving both primary and non-primary visual cortex (V2/V3; Chubb and Sperling, 1988; Larsson et al., 2006). Given that parallels have often been drawn between visual and auditory cortical functional organization (Rauschecker and Tian, 2000), we were interested in examining how well characterization of sounds by their acoustic complexity might reflect new insights into regional functional specialization.

Given that auditory neuroimaging studies exhibit a high degree of stimulus and task heterogeneity, their individual cortical activity patterns are not easily integrated to obtain an unambiguous picture of typical human auditory cortical organization. Neuroimaging meta-analysis offers a potential solution to this problem as it estimates the consistency of regional brain activity across similar stimuli and tasks, providing a quantitative summary of the state of research in a specific cognitive domain (Fox et al., 1998), estimating the replicability of effects across different scanners, tasks, stimuli, and research groups. By revealing consistently activated voxels across a set of experiments, meta-analysis can characterize the cortical response specificity associated with a particular type of task or stimulus (Wager et al., 2009). Activation Likelihood Estimation (ALE) is a voxel-wise meta-analysis method that provides a quantitative summary of task-related activity consistency across neuroimaging studies (Turkeltaub et al., 2002).

In the current study, we use quantitative ALE meta-analysis to examine the spatial consistency of human auditory processing, classified using either conventional sound categories or acoustic complexity. Given the focus of our study on stimulus complexity effects, we excluded studies of spatial auditory processes including localization, and inter-aural delay, as well as those including complex tasks.

First, we classified sounds using typical auditory *categories* to examine the evidence supporting hierarchically and hemispherically lateralized functional organization for auditory cortical processing. Hierarchical auditory processing has been described as sensitivity to stimulus complexity increasing from primary to non-primary auditory cortex, with simpler perceptual features represented at primary levels (Wessinger et al., 2001; Hall et al., 2002; Scott and Johnsrude, 2003). Relative hemispheric specialization is reflected by predominantly left-hemisphere processing for speech sounds and stronger right-hemisphere responses to music (for a review see Zatorre et al., 2002). We used typical sound *categories*, such as pure tones, noise, music, and vocal sounds, to classify auditory material to see if simple sound processing is associated with

activity in primary auditory cortex while complex sound processing is associated with activity including both primary and non-primary auditory cortex. We were also interested in examining whether there was meta-analytic evidence for distinctive patterns of hemispheric specialization for music and vocal sounds.

Next, we more closely examined vocal stimuli and a particular subcategory of vocal sounds: intelligible speech. Vocalizations constitute an ecologically central sound category that includes all sounds with a vocal quality irrespective of phonetic or lexical content. Examples include speech in various languages, non-speech affective vocalizations (e.g., laughter), and laboratory-engineered sounds, such as time-reversed speech, that exhibit distinctly vocal qualities. Vocal sounds include, but are not limited to, intelligible speech. Based on previous findings, we expected to observe bilateral superior temporal gyrus (STG) and the superior temporal sulcus (STS) activity related to vocal sounds (Belin et al., 2000, 2002; Kriegstein and Giraud, 2004), and anterior STG and STS activity on the left related to speech intelligibility (Benson et al., 2006; Uppenkamp et al., 2006).

Finally, we examined whether *acoustic complexity*, estimated from variations in time (temporal) and frequency (spectral) dimensions, represents a relevant organizing principle for functional response specificity in human auditory cortex. In terms of spectral composition, stimuli can have single or multiple frequency components. In the temporal dimension, stimuli can be characterized as unchanging or, for those containing temporal changes, having either regular or irregular changes. Using this classification, we characterized the cortical response related to each level of acoustic complexity. Then, by comparing the “multiple” to the “single” categories, independent of the temporal changes, and the “changing” to the “unchanging” categories, independent of the spectral composition, we isolated the cortical activity related to variations in the frequency and time dimensions, respectively.

MATERIALS AND METHODS

INCLUSION OF STUDIES

A preliminary list of articles was identified using several Medline database searches including both articles published prior to March 2010 [keywords: positron emission tomography (PET), functional magnetic resonance imaging (fMRI), auditory, sound, hear*, speech, and music] and lists of citations within those articles. Studies were included if they fulfilled specific inclusion criteria: (1) the study was published in a peer-reviewed journal; (2) the study involved a group of healthy typical adult participants with no history of hearing, psychiatric, neurological, or other medical disorders, (3) the subjects were not trained musicians; (4) the auditory stimuli were delivered binaurally, with no inter-aural delay because of our focus on non-spatial auditory processing; (5) the task-related activity coordinates were reported in standardized anatomical space; (6) the study used whole-brain imaging and voxel-wise analysis. As our main goal was to determine the spatial distribution within auditory cortical regions, the few studies using incomplete brain coverage, but that included the temporal cortex were not excluded (Binder et al., 1996, 2000; Belin et al., 1999; Celsis et al., 1999; Hugdahl et al., 2003; Stevens, 2004; Schönwiesner et al., 2005; Zaehle et al., 2007). Additionally, some studies specifically included subcortical structures (Griffiths et al., 1998; Hwang et al., 2007; Mutschler

et al., 2010); and (7) the study had to include passive listening or a simple response task, such as a button press at the end of each sound to assess the participants' attentive state, task characteristics that tended to minimize the inclusion of activity related to top-down processes or task difficulty (Dehaene-Lambertz et al., 2005; Dufor et al., 2007; Sabri et al., 2008).

Of over 7000 articles retrieved, 58 (19 PET and 39 fMRI) satisfied all inclusion criteria and were included in the analysis (Table 1). Several studies reported activity from multiple task and control conditions. For our analysis, only conditions incorporating either no overt task or a simple task used to maintain attention were considered. To maintain consistency among the control conditions, only task contrasts with a low-level baseline (silence, tone, or noise) were included. For some studies, more than one contrast satisfied our criteria and all were included in the analysis. This procedure was employed to maximize the sensitivity of the analysis, but could potentially bias the results toward samples for which more than one contrast was included.

CONTRAST CLASSIFICATION PROCEDURE

One hundred seventeen contrasts, including 768 foci, met the inclusion criteria. These contrasts were classified first by typical sound categories and then according to their variation along either the frequency or time dimension (Table 1).

For the first method, each contrast was classified with respect to one of the typical sound categories: simple sounds or pure tones (9 contrasts, 22 foci), noise (4 contrasts, 31 foci), music (10 contrasts, 175 foci), and vocal sounds (62 contrasts, 370 foci). The pure tones category included only contrasts of single tones vs. silence; the noise category included white, pink, and brown noise (Rimol et al., 2005), noise bursts (Zatorre et al., 1992), and the combination of multiple reversed environmental sounds (Zatorre et al., 2002). Melodies, notes, chords, and chord progressions were classified as music. Finally, all sounds with a vocal quality (syllables, words, voices, reversed words, or pseudowords) were included in the vocal sounds category. Ideally, we would have included other commonly used sound categories such as animal or environmental sounds; however the number of contrasts falling into these categories was not sufficient for quantitative meta-analysis, with only one contrast presenting environmental sounds and only two falling into the animal sound category. The remaining contrasts (30/117), including modulated tones, frequency sweep, harmonic tones, or recorded noise, were not included in this analysis because they did not neatly fit into one sound category.

For the second method, we classified the stimuli with respect to their acoustic features. Two levels of complexity were defined using the frequency dimension (single and multiple frequency components) and three levels in the time domain (unchanging, regular periodic change, or irregular change). Therefore, task contrasts were classified in one of six complexity levels depending on their frequency- and time-related acoustic features (Table 1; Figure 5A): (1) "single, unchanging" (single tone; 9 contrasts, 22 foci), (2) "single, regular change" (frequency or amplitude modulated tone, single formant frequency sweep, parametric variation of modulation rate or rate of presentation; 8 contrasts, 38 foci), (3) "single, irregular change" (1 contrast, 4 foci), (4) "multiple, unchanging" (harmonic tone, square wave tone, vowel, noise, or parametrically increasing

spectral component numbers; 10 contrasts, 57 foci), (5) "multiple, regular change" (tone sequences and increasing click rate sequences; 6 contrasts, 41 foci), or (6) "multiple, irregular change" (vocal sounds, music, or environmental sounds; 70 contrasts, 517 foci). Each task contrast was classified using the stimulus description provided in each study. Contrasts resulting from covariate effects of a parameter of interest were classified according to parameter complexity. For instance, effects related to parametric increases in temporal modulation rate were assigned to the "single, regular change" complexity level (Schönwiesner et al., 2005). Ambiguous contrasts were excluded from analysis. For example, we did not classify contrasts that used comparison stimuli that had acoustic complexity comparable to the stimuli of interest (Zatorre et al., 1994; Griffiths et al., 1998; Blood et al., 1999; Mummery et al., 1999; Warren and Griffiths, 2003; Giraud et al., 2004; Schwarzbauer et al., 2006; Peretz et al., 2009) nor those using stimuli that could be assigned to more than one complexity level, such as notes, chords, or chord progressions (i.e., stimuli including note/chord/chord progression; Benson et al., 2001).

ALE META-ANALYSIS

After the task-related activity maxima were classified, ALE maps (Turkeltaub et al., 2002) were computed using GingerALE 1.1 (Laird et al., 2005). Coordinates reported in MNI space were converted to Talairach space using the Lancaster transform *icbm2tal* (Lancaster et al., 2007). ALE models uncertainty in localization of each activation focus as a Gaussian probability distribution, yielding a statistical map in which each voxel value represents an estimate of the likelihood of activity at that location, utilizing a fixed effects model for which inferences should be limited to the studies under examination. Critical thresholds for the ALE maps were determined using a Monte Carlo style permutation analysis of sets of randomly distributed foci. A FWHM of 10 mm was selected for the estimated Gaussian probability distributions. Critical thresholds were determined using 5000 permutations, corrected for multiple comparisons ($p < 0.01$ false discovery rate, FDR; Laird et al., 2005) with a cluster extent of greater than 250 mm³. In order to present results in the format most commonly used in the current literature, the ALE coordinate results were transformed into MNI standard space using the Lancaster transform (Lancaster et al., 2007), while ALE maps were transformed by applying spatial normalization parameters obtained from mapping from Talairach to MNI space.

ANALYSIS USING CLASSIFICATION BY TYPICAL AUDITORY CATEGORIES

First, ALE maps were computed for each of the four typical auditory categories: pure tones, noise, music and vocal sounds. Each resulting map shows regions exhibiting consistent activity across studies for each sound category. For example, the "music" map shows the voxel-wise probability of activity for all "musical stimuli vs. baseline" contrasts.

Next, we examined hemispheric specialization effects by directly comparing the "music" and "vocal" sound categories. We directly compared a random subsample of the "vocal" sounds category (Table A1 in Appendix; 20 contrasts, 156 foci) to the "music" category (10 contrasts, 175 foci). This procedure ensured that the resulting ALE maps would reflect activity differences between studies rather than the imbalance in coordinate numbers

Table 1 | Neuroimaging studies included in the meta-analysis.

Reference	Table/figure	Category	Frequency composition	Changes over time	Stimuli of interest	Baseline	Foci
Belin et al. (1998)	Table 1, p. 537	–	Multiple	Irregular	Rapid formant transition	Silence	3
Belin et al. (1999)	Text, p. 422–423	–	Multiple	Irregular	Extended formant transition	Silence	2
Benson et al. (2001)	Table 1, p. 372	–	Multiple	Unchanging	Harmonic complex sound	Silence	4
Binder et al. (1996)	Table 2, p. 373	Music	–	–	Note/chord/chord progression	Rest	27
Binder et al. (2000)	Table 1, p. 1244	Vocal	Multiple	Irregular	Vowel/consonant–vowels/syllables	Rest	24
	Table 2, p. 518	Vocal	Multiple	Irregular	Words	Tones sequence	1
	Appendix	–	Multiple	Regular	Sequence of tones	Noise	11
		Vocal	Multiple	Irregular	Words	Sequence of tones	8
		Vocal	Multiple	Irregular	Pseudowords	Sequence of tones	8
		Vocal	Multiple	Irregular	Reversed words	Sequence of tones	7
Binder et al. (2008)		Vocal	Multiple	Irregular	Words	Rest	12
		Vocal	Multiple	Irregular	Words	Tones	2
Blood et al. (1999)	Table 2, p. 384	Music	–	–	Melodies	Noise	3
		Music	–	–	Melodies	Noise	3
Brown et al. (2004)	Table 1, p. 2035	Music	Multiple	Irregular	Music	Rest	21
Burton et al. (2000)	Table 2, p. 682	Vocal	Multiple	Irregular	Words	Tone	3
		Vocal	Multiple	Irregular	Words	Tone	6
Burton and Small (2006)	Table 2, p. 647	Vocal	Multiple	Irregular	Words	Tone	4
Celsis et al. (1999)	Table 1, p. 138	Tone	Single	Unchanging	Tone	Rest	1
		Tone	Single	Unchanging	Tone	Rest	1
		Vocal	Multiple	Irregular	Syllable	Rest	3
		Tone	Single	Unchanging	Tone	Rest	3
		–	Multiple	Unchanging	Square wave tone	Rest	3
		Vocal	Multiple	Irregular	Syllable	Rest	3
Chen et al. (2008)	Table 2, p. 2849	Music	Multiple	Irregular	Musical rhythms	Silence	9
Dalla Barba et al. (1998)	Table 1, p. 552	Vocal	Multiple	Irregular	Words	Rest	5
De Nil et al. (2008)	Table 2, p. 119	Vocal	Multiple	Irregular	Words	Silence	1
Démonet et al. (1992)	Table 4, p. 1758	Vocal	Multiple	Irregular	Phonemes (three or four syllables non-words)	Tones	7
		Vocal	Multiple	Irregular	Words	Tones	14
Engelien et al. (2006)	Table 1, p. 603	–	Multiple	Irregular	Meaningful non-verbal sounds	Rest	8
Fernandez-Espejo et al. (2008)	Table 2, p. 887	–	Multiple	Irregular	Meaningless non-verbal sounds	Rest	14
Gaab et al. (2007a)	Table 1, p. 710	Vocal	Multiple	Irregular	Narrative sentences	Silence	2
		Vocal	Multiple	Irregular	Words	Silence	7
		Vocal	Multiple	Irregular	Words	Silence	9
Gaab et al. (2007b)	Table 1, p. 727	Vocal	Multiple	Irregular	Words	Silence	5
		–	Multiple	Irregular	Recorded scammer noise	Silence	2
Giraud et al. (2004)	Table 1, p. 250	Vocal	–	–	Speech (sentences)	Temporally matched complex noise	5
Griffiths et al. (1998)	Table 1, p. 424	Music	–	–	Melodies	Sequence of tones	4
Hall et al. (2000)	Table 1, p. 114	–	Single	Regular	Amplitude and frequency modulated tone	Static tone	2
Hall et al. (2002)	Table 2, p. 144	–	Multiple	Unchanging	Harmonic complex tone	Single tone	5

Hart et al. (2003)	Text, p. 778	–	Single	Regular	Frequency modulated tone	Static tone	10
Hart et al. (2004)	Table 1, p. 182	Tone	Single	Regular	Amplitude and frequency modulated tone	Unmodulated tone	2
Hertrich et al. (2010)	Table S1	Vocal	Multiple	Irregular	Frequency modulated tone	Silence	4
Hugdahl et al. (2003)	Table 3, p. 43	–	Single	Regular	Syllables	Silence	6
	Table 4, p. 43	Vocal	Multiple	Irregular	Single formants sweep	Silence	10
	Table 5, p. 43	Vocal	Multiple	Irregular	Vowel	Silence	8
Hwang et al. (2007)	Table 2, p. 289	Vocal	Multiple	Irregular	Pseudowords	Silence	4
Jäncke et al. (1998)	Table 1, p. 878	Vocal	Multiple	Irregular	Words	Silence	3
	Table 2, p. 878	–	Multiple	Irregular	Sentences	Silence	2
Lillywhite et al. (2010)	Table 1, p. 875	Vocal	Multiple	Irregular	Syllables	Quiet	7
Müller et al. (1999)	Table 3, p. 24	–	Multiple	Irregular	Syllables	Silence	9
Mummers et al. (1999)	Table 1, p. 452	Vocal	Multiple	Irregular	Syllables	Silence	9
Mutschler et al. (2010)	Supp. Material	Vocal	Multiple	Regular	Tone sequence	Silence	4
O'Leary et al. (1996)	Table 2, p. 27	Vocal	Multiple	Irregular	Tone sequence	Silence	6
	Table 1, p. 452	Vocal	Multiple	Irregular	Tone sequence	Silence	6
Mummers et al. (1999)	Table 1, p. 452	Vocal	Multiple	Irregular	Story listening	White noise	2
Mutschler et al. (2010)	Supp. Material	Music	Multiple	Regular	Sequence of tones	Rest	6
O'Leary et al. (1996)	Table 2, p. 27	Vocal	Multiple	Irregular	Sentences	Rest	5
	Table 1, p. 452	Vocal	Multiple	–	Speech (words)	Signal correlated noise	7
Mutschler et al. (2010)	Supp. Material	Music	Multiple	Irregular	Melodies	Rest	94
O'Leary et al. (1996)	Table 2, p. 27	Vocal	Multiple	Irregular	Words	Tone	2
	Table 1, p. 452	Vocal	Multiple	Irregular	Environmental sounds	Tone	2
Obleser et al. (2006)	Table 2, p. 566	–	Multiple	Irregular	Vowel	Band-pass noise	4
Pauls et al. (1995)	Table 2, p. 667	Vocal	Multiple	Irregular	Words	Tones	5
Peretz et al. (2009)	Table 1, p. 263	Music	–	–	Melodies	Random tone sequence	8
Petersen et al. (1988)	Table 2, p. 585	Vocal	Multiple	Irregular	Words	Silence (fixation point)	6
Price et al. (1992)	Figure 1, p. 180	Vocal	Multiple	Irregular	Nouns	Rest	6
Reynolds Losin et al. (2009)	Table 1, p. 374	Vocal	Multiple	Irregular	Story listening	Rest	3
Rimol et al. (2005)	Table 1, p. 1063	Vocal	Multiple	Irregular	Consonant	Rest	8
	Table 2, p. 1064	Vocal	Multiple	Irregular	Syllables	Rest	6
Schwarzbauer et al. (2006)	Table 1, p. 779	Vocal	–	–	Noise (white, brown, pink)	Rest	12
	Table 1, p. 779	Vocal	–	–	Noise (white, brown, pink)	Rest	8
Schönwiesner et al. (2005)	Text, p. 1523–1525	–	–	–	Consonant	Noise	1
	Table 1, p. 1950	Tone	–	–	Syllable	Noise	3
Specht and Reul (2003)	Table 1, p. 1950	Tone	–	–	Sentences	Signal correlated noise	12
	Table 1, p. 1950	Tone	–	–	Sentences	Signal correlated noise	6
	Table 1, p. 1950	Tone	–	–	Parametric increase in temporal rate change	Silence	2
	Table 1, p. 1950	Tone	–	–	Parametric increase in spectral components	Silence	2
	Table 1, p. 1950	Tone	–	–	Tones	Rest	5

(Continued)

Table 1 | Continued

Reference	Table/figure	Category	Frequency composition	Changes over time	Stimuli of interest	Baseline	Foci
		-	Multiple	Irregular	Sounds of animal and instruments	Rest	9
		Vocal	Multiple	Irregular	Words	Rest	18
	Table 2, p. 1951	-	Multiple	Irregular	Sounds of animal and instruments	Tones	7
		Vocal	Multiple	Irregular	Words	Tones	12
Steinbrink et al. (2009)	Table 1, p. 2406	Vocal	Multiple	Irregular	Increase rate of syllable presentation	Silence	7
		-	Multiple	Regular	Increase rate of click sequence	Silence	8
Stevens (2004)	Table 1, p. 166	Vocal	Multiple	Irregular	Voice	Tone	13
		Vocal	Multiple	Irregular	Words	Tone	5
Thivard et al. (2000)	Table 1, p. 2971	-	Multiple	Unchanging	Static complex sound	Silence	6
		-	Multiple	Irregular	Complex sound modulated (motion)	Silence	6
		-	Single	Irregular	Complex sound modulated (motion)	Static complex sound	4
Vouloumanos et al. (2001)	Table 1, p. 998	Vocal	Multiple	Irregular	Words	Tones	8
		-	Multiple	Irregular	Non-speech complex sinewave sounds	Tones	5
Warren and Griffiths (2003)	Table 1, p. 5802	Music	-	-	Changing pitch sequences	Fixed pitch sequences	4
Warriner et al. (2009)	Table 2, p. 65	-	Single	Regular	Parametric increase in temporal rate change	Silence	6
		-	Multiple	Unchanging	Parametric increase in spectral components	Silence	6
Wise et al. (1991)	Table 2, p. 1808	Vocal	Multiple	Irregular	Nouns	Rest	6
Zaehle et al. (2007)	Table 1, p. 1201	Vocal	Multiple	Irregular	Sentences	Silence	2
		Vocal	Multiple	Irregular	Sentences	Silence	2
		Tone	Single	Unchanging	Tones	Silence	2
		Tone	Single	Unchanging	Tones	Silence	2
		Vocal	Multiple	Irregular	Sentences	Tones	2
		Vocal	Multiple	Irregular	Sentences	Tones	2
		Vocal	Multiple	Irregular	Sentences	Silence	2
		Vocal	Multiple	Irregular	Sentences	Silence	2
	Table 2, p. 1201	Vocal	Multiple	Irregular	Sentences	Silence	2
		Tone	Single	Unchanging	Tones	Silence	2
		Tone	Single	Unchanging	Tones	Silence	2
		Vocal	Multiple	Irregular	Sentences	Tones	2
		Vocal	Multiple	Irregular	Sentences	Tones	2
Zatorre et al. (1992)	Table 2, p. 847	Noise	Multiple	Unchanging	White noise bursts	Silence	8
		Vocal	Multiple	Irregular	Syllables	Noise	6
Zatorre et al. (1994)	Table 2, p. 1911	Music	-	-	Melodies	Noise	2
Zatorre et al. (1996)	Table 6, p. 26	Vocal	Multiple	Irregular	Syllables	Noise	6
Zatorre et al. (2002)	Table 1, p. 907	Noise	Multiple	Irregular	Noise (reversed environmental sounds)	Silence	3
Zatorre and Belin (2001)	Table 1, p. 948	-	Single	Regular	Tone sequence with increasing rate of presentation	Silence	2
		-	Multiple	Unchanging	Tone sequence with increasing number of spectral elements	Silence	3
Zevin and McCandliss (2005)	Table 1	Vocal	Multiple	Irregular	Syllables	Silence	8

between those categories (Laird et al., 2005). Then, as lateralization effects are reported for intelligible speech rather than vocal sounds, only contrasts using intelligible speech with semantic content, such as words or sentences, were included. The “music” and the “speech” categories were directly compared to investigate the expected lateralization effects. Given that many contrasts fell into the intelligible speech category, we selected only one contrast per study (Table A1 in Appendix), including a total of 27 contrasts (166 foci).

Finally, we assessed cortical auditory specialization for processing intelligible speech. Given that specialized auditory processes can be more easily isolated when the contrasting stimuli are as close as possible to the stimuli of interest in terms of acoustic complexity (Binder et al., 2000; Uppenkamp et al., 2006), contrasts containing unintelligible spectrally and temporally complex sounds were used as for comparison. Thirteen contrasts (76 foci, see Table A1 in Appendix) selected included reversed words, pseudowords, recorded scanner noise, single formant, environmental sounds, and modulated complex sounds. We directly compared the intelligible speech and complex non-speech sound categories.

ANALYSIS USING CLASSIFICATION BY AUDITORY COMPLEXITY

To investigate the relevance of acoustic complexity as a stimulus property predicting functional auditory specialization, we computed ALE maps for each level of complexity. Given that only one contrast fell into the “single, irregular change” dimension, this analysis was not conducted. Moreover, as most of the contrasts were classified as “multiple frequencies, irregular modulation” (70 contrasts, 517

foci), a randomly selected subsample of 10 contrasts (70 foci, see Table A1 in Appendix) were selected from this level of complexity to facilitate comparison of activity extent between levels.

Next, we examined effects related to auditory complexity. For the frequency domain, all contrasts falling in the “multiple” level (26 contrasts, 168 foci) were directly compared to those in the “single” level (18 contrasts, 64 foci), independent of their variation over time, (Figure 5A, bottom row vs. top row, green arrow). For the time dimension, comparisons were made between the contrasts including stimulus changes over time (regular and irregular; 25 experiments, 153 foci) and those who did not (unchanging; 19 contrasts, 79 foci), independent of their frequency composition (Figure 5A, middle and right column vs. left column, blue arrow).

RESULTS

STIMULUS CLASSIFICATION USING TYPICAL AUDITORY CATEGORIES

We observed different patterns of activity corresponding to the typical sound categories of pure tones, noise, music, and vocal sounds (Figure 1; Table 2). For all the categories, the strongest effects were found in auditory cortex (Brodmann areas 41, 42, and 22). For the pure tone map, high ALE values were found bilaterally in medial Heschl’s gyri (HG). The noise map revealed effects in right medial HG and bilaterally in STG posterior and lateral to HG. Effects related to music were seen in HG, anterior and posterior STG. Finally, vocal sounds elicited large bilateral clusters of activity in HG as well as anterior, posterior, and lateral aspects of the STG. While pure tone effects were restricted to auditory cortex, effects outside temporal cortex were observed for the other categories. Additional activity was seen in frontal cortex for noise (BA 6, 9),



FIGURE 1 | Activation Likelihood Estimation maps showing clusters of activity related to sound categories: pure tones, noise, music, and vocal sounds.

Maps are superimposed on an anatomical template in MNI space. Axial images are shown using the neurological convention with MNI z-coordinate labels ($p_{FDR} < 0.01$).

Table 2 | Category classification.

Region	BA	Coordinates			Volume (mm ³)	ALE (×10 ⁻³)
		x	y	z		
PURE TONES						
Temporal						
Heschl's gyrus	41	53	-17	1	3864	26.00
Heschl's gyrus	41	-51	-20	1	2600	16.85
NOISE						
Temporal						
Superior temporal gyrus	22	68	-23	5	1544	12.98
Superior temporal gyrus	42	-50	-29	15	1384	15.58
Superior temporal gyrus	22	51	-24	-1	648	11.30
Frontal						
Precentral gyrus	9	-46	12	38	664	8.67
Superior frontal gyrus	6	13	14	57	552	11.36
Inferior frontal gyrus	9	45	14	20	496	10.28
Sub cortical						
Cerebellum		33	-65	-28	360	7.83
MUSIC						
Temporal lobe						
Superior temporal gyrus	22	53	-11	2	6744	21.07
Superior temporal gyrus	22	-51	-12	-7	4848	28.52
Superior temporal gyrus	22	55	12	-9	1096	18.21
Superior temporal gyrus	22	-49	3	-13	256	11.58
Frontal lobe						
Inferior frontal gyrus	45	51	34	18	1264	14.91
Inferior frontal gyrus	45	-42	24	12	1240	14.04
Precentral gyrus	44	-57	17	4	360	12.41
Medial frontal gyrus	6	-2	-3	67	352	14.39
Precentral gyrus	4	54	-2	51	280	12.31
Sub cortical						
Cerebellum		-31	-62	-20	912	14.10
VOCAL SOUNDS						
Temporal						
Superior temporal gyrus	22	-59	-12	-5	22648	134.64
Superior temporal gyrus	22	62	-13	-4	18088	101.69
Frontal						
Inferior frontal gyrus	45	-48	17	26	2112	21.16
Limbic						
Cingulate gyrus	24	-5	13	46	584	18.72

MNI coordinates of the locations of significant ALE maxima for each sound category ($p < 0.01$, $k = 250$ voxels).

music (BA 4, 6, 44, 45, 46), and vocal sounds (BA 45). Effects were observed in cerebellum for noise and music as well as in the anterior cingulate gyrus for vocal sounds.

Effects related to typical sound categories were lateralized. Qualitative examination revealed larger clusters in right auditory cortex for music and in left auditory cortex for vocal sounds (Table 2). The direct comparisons between the musical and vocal sounds and between the musical and speech sounds yielded similar findings (Figure 2; Table 3). Greater activity related to music was observed bilaterally in posterior and anterolateral HG, the planum polare, and the most anterior parts of the right STG. We also observed effects

related to music processing outside the temporal lobe, in inferior frontal gyrus (BA 45), the middle frontal gyrus (BA 6), and the left cerebellum (lobule IV). On the other hand, the reverse comparisons revealed stronger activity for vocal sounds as well as for speech in lateral HG, extending to lateral and anterior STG. For the vocal sounds, the extent of auditory activity was greater on the left (10312 voxels) than on the right (4952 voxels), however the ALE values were similar on the left (45.66×10^{-3}) and on the right (42.24×10^{-3}). As for the speech sounds, both the volume of activity and the corresponding ALE were greater on the left (11112 voxels, 61.39×10^{-3}) than the right (5736 voxels, 38.21×10^{-3}) hemisphere.

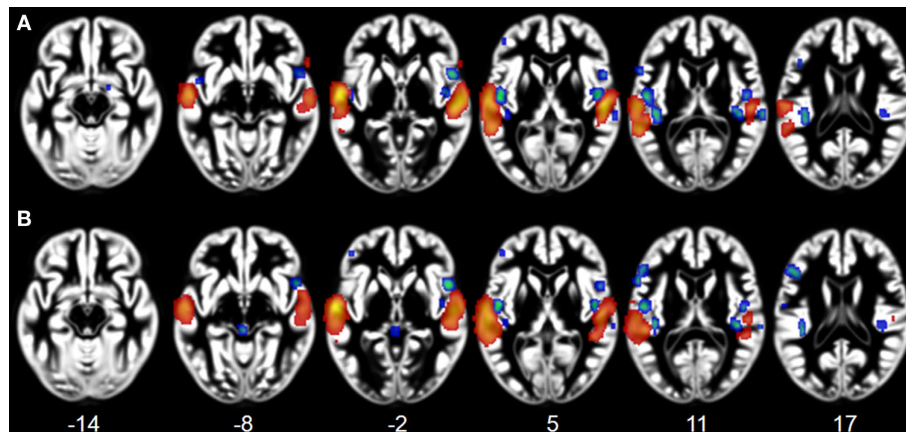


FIGURE 2 | Activation Likelihood Estimation maps showing lateralization effects for (A) voices > music (RED–YELLOW) and music > voices (BLUE–GREEN) comparisons and for (B) speech > music (RED–YELLOW) and music > speech (BLUE–GREEN) comparisons. Maps are superimposed on an anatomical template in MNI space. Axial images are shown using the neurological convention with MNI z-coordinate labels ($p_{FDR} < 0.01$).

We observed specialization for speech processing in auditory cortex. The comparison between intelligible speech and complex non-speech sounds, including vocal sounds without intelligible content, is shown in **Table 4** and **Figure 3A**. Speech was associated with greater activity in non-primary (BA 22) and associative (BA 39) auditory areas, lateral STG, bilateral anterior and middle STS, and the planum temporale (PT). These clusters were larger and had higher ALE values in the left hemisphere. We also observed stronger left prefrontal cortical activity (BA 8) for speech sounds. The reverse comparison yielded stronger activity related to complex non-speech sounds in the right PT ($x = 68, y = -27, z = 8, 128$ voxels; **Figure 3A**). The ALE maps associated with speech intelligibility had overlap with the vocal sound category maps (**Figure 3B**). While large bilateral clusters were observed along the STG and STS for the vocal sounds, there was specific sensitivity to speech intelligibility in the left anterior STG.

STIMULUS CLASSIFICATION USING AUDITORY COMPLEXITY

Classification of sounds with respect to their spectral and temporal complexity revealed effects in the temporal lobe (**Table 5**; **Figure 4**). The “single, unchanging” stimulus class was associated with two clusters centered on medial HG (BA 41). The “single, regular change” stimulus class was associated with two large bilateral clusters of activity in medial and lateral HG, extending around HG into the anterolateral STG. On the left, we observed one additional peak of activity in posterior STG. For the “multiple, unchanging” stimulus class, temporal lobe activity was centered on medial HG and posterior STG. Effects for the “multiple, regular change” stimulus class were observed in HG, extending to the posterolateral STG. Finally, the “multiple, irregular change” stimulus class was associated with large bilateral effects in, and posterior to, HG. The complexity level maps revealed effects outside the temporal lobe, in frontal cortex areas BA 6, 9, 36, and 47 for the “multiple, unchanging” and “multiple, regular change” stimulus classes. We also observed effects in the cerebellum for the “single, regular change” and “multiple, irregular change” stimulus classes.

Effects related to stimulus spectral and temporal variations were identified by comparing, respectively, the multiple to the single stimulus classes (independent of changes over time; **Figure 5B**,

GREEN) and the changing to the unchanging stimulus classes (independent of the number of frequency components; **Figure 5B**, BLUE). The coordinates of the effects related to increasing auditory complexity are reported in **Table 5**. Overlapping sensitivity to spectral and temporal effects was observed in the lateral portion of HG. Increasing numbers of frequency components were associated with greater effects in posterior and lateral non-primary auditory fields, specifically bilateral posterolateral STG and PT. Modulatory effects were also seen in inferior frontal gyrus (BA 45, 47). In contrast, the effects related to temporal modulations compared to their absence were observed in HG, anterior STG, anterior STS, inferior frontal cortex (BA 46, 47), and right cerebellum (lobule IV).

DISCUSSION

SUMMARY OF FINDINGS

In a quantitative meta-analysis of 58 neuroimaging studies, we examined the functional specialization of human auditory cortex using two different strategies for classifying sounds. The first strategy employed typical categories, such as pure tones, noise, music, and vocal sounds. The second strategy categorized sounds according to their acoustical (spectral and temporal) complexity.

Activation Likelihood Estimation maps computed for each typical sound category included simple (pure tones) and complex (noise, voices, and music) sounds. This analysis gave results consistent with models describing hierarchical functional organization of the human auditory cortex, with simple sounds eliciting activity in the primary auditory cortex and complex sound processing engaging additional activity in non-primary fields. We observed an expected leftward hemispheric specialization for intelligible speech, while right-hemisphere specialization for music was less evident. Additionally, the comparison of intelligible speech to complex non-speech stimuli yielded bilateral effects along the STG and STS, with higher sensitivity to speech intelligibility in the left anterior STG.

Examining an alternative classification based on stimulus variation along spectral and temporal dimensions, we observed a within-hemisphere functional segregation, with spectral effects strongest in posterior STG and temporal modulations strongest in anterior

Table 3 | Lateralization effects.

Region	BA	Coordinates			Volume (mm ³)	ALE ($\times 10^{-3}$)
		x	y	z		
MUSIC > VOICES						
Temporal						
Heschl's gyrus	41	-49	-11	4	920	22.64
Heschl's gyrus	41	-40	-30	10	800	15.25
Heschl's gyrus	41	51	-9	0	680	17.15
Superior temporal gyrus	22	53	12	-9	584	16.18
Heschl's gyrus	41	42	-27	11	336	13.18
Frontal						
Inferior frontal gyrus	45	51	34	18	776	14.75
Inferior frontal gyrus	45	-40	27	14	384	13.00
Middle frontal gyrus	6	-2	-3	67	376	14.39
Subcortical						
Cerebellum		-31	-65	-20	392	13.04
VOICES > MUSIC						
Temporal						
Superior temporal gyrus	22	-59	-10	-5	10312	45.66
Superior temporal gyrus	22	59	-19	-1	4952	42.24
Middle frontal gyrus	9	-48	19	26	1032	17.67
MUSIC > SPEECH						
Temporal						
Superior temporal gyrus	22	55	12	-9	728	17.31
Heschl's gyrus	41	-49	-9	4	696	20.77
Heschl's gyrus	41	-38	-28	8	480	12.75
Heschl's gyrus	41	51	-9	2	408	14.44
Heschl's gyrus	41	45	-27	11	336	15.08
Frontal						
Inferior frontal gyrus	45	-40	27	14	1048	13.32
Inferior frontal gyrus	45	51	34	18	560	14.37
Middle frontal gyrus	6	-2	-5	68	304	14.33
Subcortical						
Cerebellum		-31	-62	-20	728	14.09
SPEECH > MUSIC						
Temporal						
Superior temporal gyrus	22	-59	-12	-5	11112	61.39
Superior temporal gyrus	22	62	-13	-4	5736	38.21
Frontal						
Medial frontal gyrus	8	-16	37	37	336	14.13

MNI coordinates of the locations of significant ALE maxima resulting from the comparison between speech and music ($p < 0.01$, $k = 250$ voxels).

temporal STG. We suggest that acoustic complexity might represent a valid alternative classificatory scheme to describe a novel within-hemisphere dichotomy regarding the functional organization for auditory processing in temporal cortex.

HIERARCHICALLY AND HEMISPHERICALLY SPECIALIZED ARCHITECTURES FOR AUDITORY PROCESSING

Originally elaborated on the basis of non-human primate studies, the hierarchical functional organization scheme in auditory cortex incorporates three levels of processing: core (primary area), belt and parabelt (non-primary areas). Simple sound processing is thought to solely recruit the core region whereas complex sounds are believed

to elicit activity in core, belt, and parabelt areas. While belt region responses are thought to be sensitive to acoustic feature variations, the parabelt, and more anterior temporal regions, show greater sensitivity to complex sounds such as vocalizations (Rauschecker, 1998; Hackett, 2008; Rauschecker and Scott, 2009; Woods and Alain, 2009). Our quantitative meta-analysis using typical sound classes confirmed that hierarchical processing is a feature that can adequately describe human auditory cortical organization.

Using an ALE analysis of pure tone processing to investigate the correspondence between the core region and activity related to simple sound processing, we observed ALE extrema values bilaterally in medial HG, the putative location of primary auditory

Table 4 | Functional specialization for speech sounds.

Region	BA	Coordinates			Volume (mm ³)	ALE (×10 ⁻³)
		x	y	z		
INTELLIGIBLE SPEECH > COMPLEX NON-SPEECH						
Temporal						
Superior temporal gyrus	22	-59	-10	-5	8760	39.48
Superior temporal gyrus	22	59	-11	-5	4416	28.91
Superior temporal gyrus	39	-48	-53	37	272	13.40
Frontal						
Medial frontal gyrus	8	-16	37	37	424	14.21
COMPLEX NON-SPEECH > INTELLIGIBLE SPEECH						
No suprathreshold voxels						

MNI coordinates of the locations of significant ALE maxima resulting from the comparison between speech and complex non-speech contrasts ($p < 0.01$, $k = 250$ voxels).

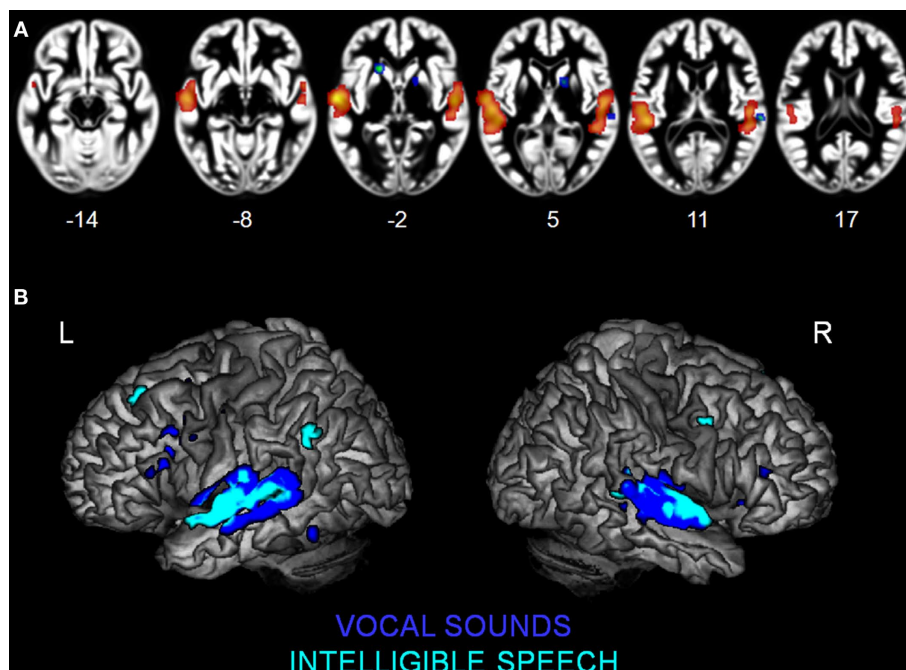


FIGURE 3 | Activation Likelihood Estimation maps showing clusters of activity related to (A) intelligible speech > complex non-speech sounds (RED-YELLOW) and to intelligible speech < complex non-speech (BLUE-GREEN). Axial images are shown using the neurological convention with

MNI z-coordinate labels. (B) Rendering of ALE maps related to vocal sound category (dark blue) and to speech intelligibility (pale blue). The maps are superimposed on anatomical templates in MNI space ($p_{FDR} < 0.01$).

cortex. This finding is consistent with previous electrophysiological (Hackett et al., 2001), cytoarchitectural (Sweet et al., 2005), and functional imaging (Lauter et al., 1985; Bilecen et al., 1998; Lockwood et al., 1999; Wessinger et al., 2001) studies of the human auditory cortex that have localized the core region to medial HG. Our findings confirm the existence of functional specialization for simple sound processing in the human core homolog. Consequently, the statistical probability maps obtained here could serve to functionally define primary auditory cortex in a region of interest analysis of functional neuroimaging data.

In contrast, we expected ALE analyses of the complex sound categories to show activity in all three levels of the processing hierarchy. We observed overlapping activity among the complex sound maps in medial HG (core) as well as stronger activity related to complex sound processing in regions surrounding medial HG, corresponding to the areas described as the auditory belt/parabelt in primates (Rauschecker, 1998; Kaas and Hackett, 2000; Rauschecker and Scott, 2009; Recanzone and Cohen, 2010) and humans (Rivier and Clarke, 1997; Wallace et al., 2002; Sweet et al., 2005). The fact that the complex sound maps showed effects in medial HG activity

Table 5 | Complexity classification.

Region	BA	Coordinates			Volume (mm ³)	ALE ($\times 10^{-3}$)
		x	y	z		
COMPLEXITY LEVELS						
Single unchanging						
Temporal						
Heschl's gyrus	41	53	-17	1	3880	26.00
Heschl's gyrus	41	-51	-18	0	2488	16.85
Single regular change						
Temporal						
Superior temporal gyrus	22	57	-11	2	6200	31.94
Superior temporal gyrus	22	-57	-15	5	4080	24.31
Sub cortical						
Cerebellum		23	-62	-21	352	7.98
Multiple unchanging						
Temporal						
Heschl's gyrus	41	-40	-27	10	7240	18.91
Heschl's gyrus	41	53	-21	1	6104	17.69
Frontal						
Inferior frontal gyrus	9	-46	12	38	480	8.67
Superior frontal gyrus	6	13	14	57	448	11.36
Inferior frontal gyrus	9	45	15	20	408	10.28
Multiple regular change						
Temporal						
Superior temporal gyrus	42	66	-21	5	6256	22.56
Superior temporal gyrus	42	-59	-21	12	6096	21.90
Frontal						
Inferior frontal gyrus	46	62	35	7	848	15.86
Inferior frontal gyrus	47	-42	27	-4	688	12.63
Multiple irregular change						
Temporal						
Superior temporal gyrus	22	57	-19	1	8344	24.12
Superior temporal sulcus	22	-57	-14	-2	6040	23.61
Subcortical						
Cerebellum		18	-58	-24	672	12.44
		-31	-64	-18	552	10.27
COMPARISONS BETWEEN COMPLEXITY LEVELS						
Multiple > single						
Temporal						
Superior temporal gyrus	22	-59	-23	10	8352	33.26
Superior temporal gyrus	22	68	-21	5	8256	42.55
Frontal						
Inferior frontal gyrus	46	62	35	7	448	15.86
Inferior frontal gyrus	47	-42	27	-4	384	13.50
Subcortical						
Cerebellum		29	-56	-26	872	14.98
Thalamus		14	-19	2	352	12.62
Changing > unchanging						
Temporal						
Superior temporal gyrus	22	59	-11	2	8808	29.96
Superior temporal gyrus	22	-57	-9	-3	5552	33.22
Superior temporal sulcus	22	-64	-34	7	336	10.54
Frontal						
Inferior frontal gyrus	46	62	35	7	440	15.85
Inferior frontal gyrus	47	-42	27	-4	360	13.50
Subcortical						
Cerebellum		23	-58	-24	1072	17.22

MNI coordinates of the locations of significant ALE maxima for each level of acoustic complexity and comparison between levels ($p < 0.01$, $k = 250$ voxels).

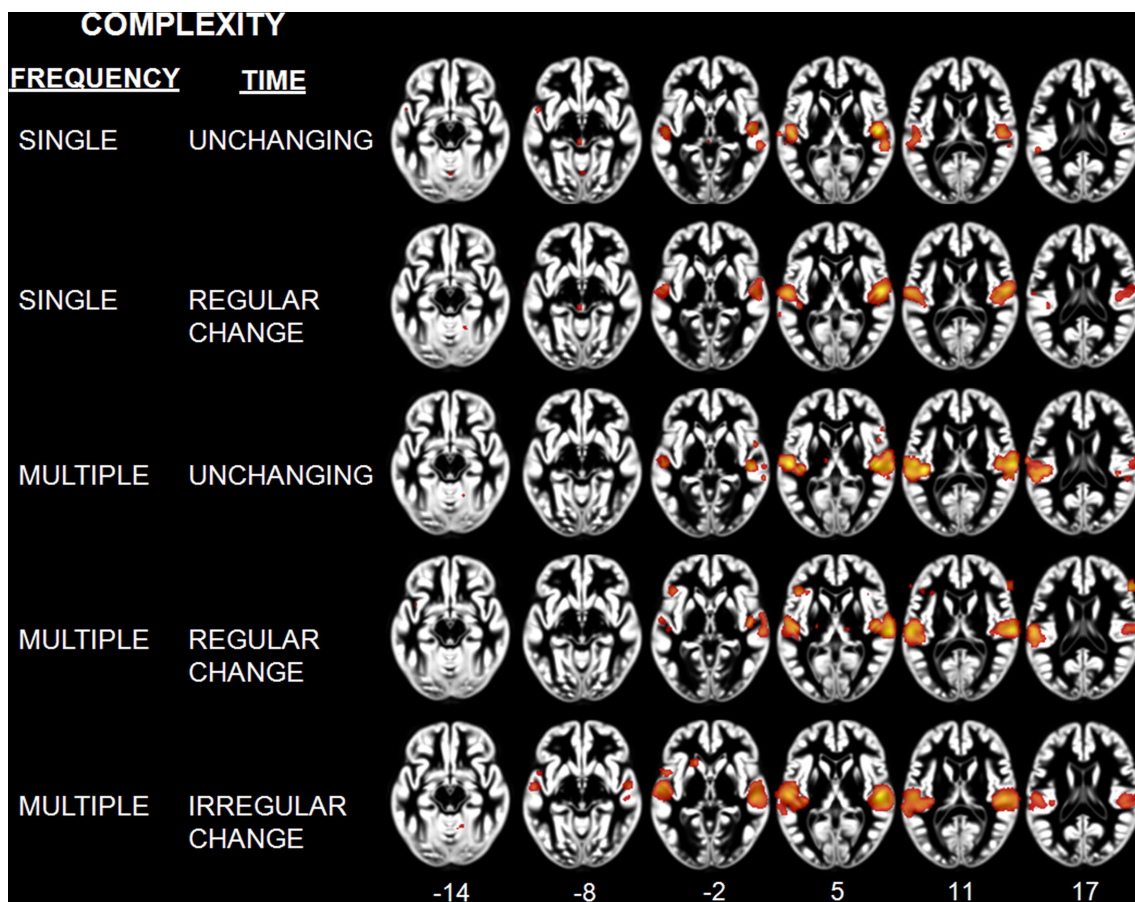


FIGURE 4 | Activation Likelihood Estimation maps showing effects related to each level of complexity: Single unchanging, single regular change, multiple unchanging, multiple regular change, and multiple irregular change. Maps are superimposed on an anatomical template in MNI space. Axial images are shown using the neurological convention with MNI z-coordinate labels ($p_{FDR} < 0.01$).

supports the notion that primary auditory regions participate in the early stages of processing upon which further complex processing is built.

Outside primary auditory cortex, noise elicited activity in posterior temporal non-primary fields such as PT. The spatial pattern was similar to that observed in relation to broadband noise, stimuli that have been used to demonstrate the hierarchical organization of human auditory cortex (Wessinger et al., 2001). The PT is generally believed to be involved in complex sound analysis and participate in both language and other cognitive functions (Griffiths and Warren, 2002).

For music, in addition to primary auditory cortex activity, we observed activity in non-primary auditory fields along the STG bilaterally. This result is consistent with the idea that simple extraction and low-level ordering of pitch information involves processes within primary auditory fields, while higher-level processing for tone patterns and melodies involve non-primary auditory fields and association cortex (Zatorre et al., 2002). Moreover, non-primary regions in anterior and posterior STG are thought to process melody pitch intervals (Patterson et al., 2002; Tramo et al., 2002; Warren and Griffiths, 2003). Music also elicited strong inferior frontal cortex activity, a region thought to process musical syntax (Zatorre et al., 1994; Maess et al., 2001; Koelsch et al., 2002).

For vocal sounds, we observed strong bilateral temporal lobe activity in anterior and posterior parts of dorsal STG and the STS, findings consistent with earlier studies (Binder et al., 1994; Belin, 2006). STG activity in response to vocal sounds has previously been interpreted as a neural correlate of the rapid and efficient processing of the complex frequency patterns and temporal variations characterizing speech. The human STG is thought to subserve complex auditory processing, such as vocalizations, as is the STG in non-human primates (Rauschecker et al., 1995). Belin and colleagues (Belin et al., 2000, 2002; Fecteau et al., 2004) reported cortical responses to voices along the upper bank of the middle and anterior STS. The anterior STS is selectively responsive to human vocal sounds (Belin et al., 2000). Response specificity to vocal sounds and their rich identity and affective information content is of crucial importance, as it reflects a set of high-level auditory cognitive abilities that can be directly compared between human and non-human primates. The regions described as “Temporal Voice Areas” in humans (Belin et al., 2000) are thought to be functionally homologous to the temporal voice regions recently described in macaques (Petkov et al., 2008). Our meta-analysis using typical sound categories demonstrates that, in humans, simple sound processing elicits activity limited to the core area while complex sounds elicit effects in all three cortical processing levels.

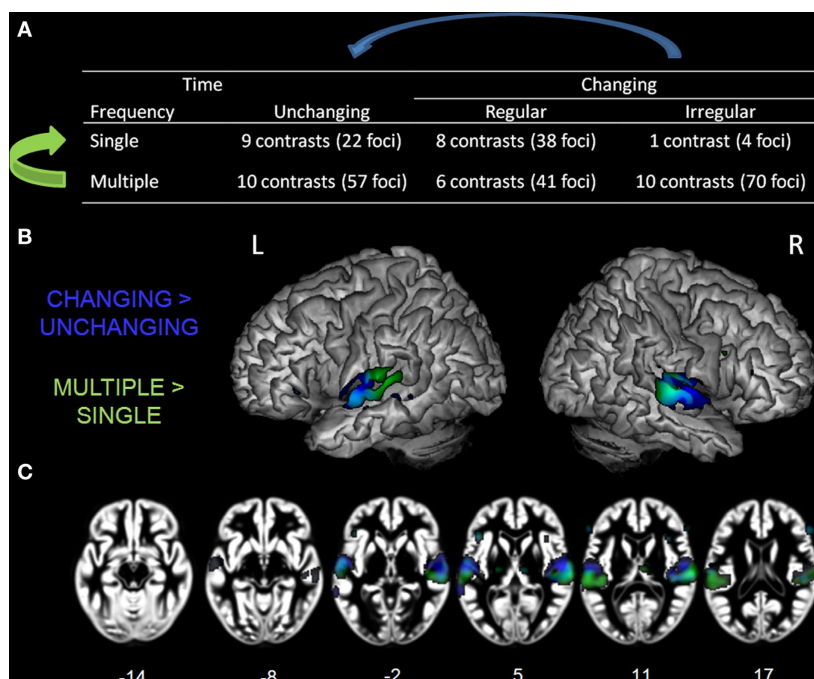


FIGURE 5 | Table of complexity levels and corresponding number of contrasts. (A) Rendering (B) and axial overlay (C) of the ALE maps reflecting the effects related to frequency (GREEN) and time (BLUE) complexity axis. Maps are superimposed on an anatomical template in MNI space. Axial images are shown using the neurological convention with MNI z-coordinate labels ($p_{FDR} < 0.01$).

In addition to the hierarchical organization of auditory cortex, we expected hemispheric asymmetries for music and speech, and observed the expected left lateralization of auditory cortex responses to vocal sounds and intelligible speech. For vocal sounds, lateralization effects were observed only as a larger volume of auditory activity on the left while, for the speech sounds, the left auditory cortical responses were larger and stronger (higher ALE values) than the right-hemisphere responses. Greater lateralization effects for intelligible speech is in agreement with previous independent imaging studies, not included in this meta-analysis, reporting that intelligible speech sounds elicit strong activity in left STG and STS (e.g., Scott et al., 2000; Liebenthal et al., 2005; Obleser et al., 2007). Conversely, we did not see the expected right response lateralization related to music. Possibly, the small number of experiments included in the music category limited the power of this analysis and could have prevented us from observing the expected rightward auditory response. ALE maps derived from small samples are more sensitive to between-study cohort heterogeneity that could limit the detection of hemispheric effects. It is also possible that the right hemisphere is sensitive to particular features of musical stimuli such as fine pitch changes (Hyde et al., 2008) or to specific task demands like contextual pitch judgment (i.e., contextual pitch judgment Warrier and Zatorre, 2004) which were not present in our sample.

RESPONSE SPECIFICITY TO SPEECH INTELLIGIBILITY

Within the general category of vocal sounds, a human-specific category of intelligible speech can be further distinguished. Response specificity to speech intelligibility is an important part

of understanding the human-specific neural network underlying speech comprehension, and ultimately human language and communication.

In order to identify speech-specific processes, we directly compared intelligible speech to complex non-speech contrasts that included unintelligible spectro-temporally complex sounds. This comparison yielded stronger speech-related activity in lateral non-primary superior temporal regions, specifically in posterior STG, and anterior and middle STS. The effects were stronger and larger in the left hemisphere. Similar effects have been reported in independent studies examining specialization for processing speech sound that did not fulfill our inclusion criteria for this analysis (Scott et al., 2000; Davis and Johnsrude, 2003; Narain et al., 2003; Thierry et al., 2003; Liebenthal et al., 2005). Consistent with the present finding, these previous reports emphasized that speech-specific STS responses are more left-lateralized.

Beyond the auditory cortex, we observed activity in left inferior frontal and prefrontal cortex. These findings support an expanded hierarchical model of speech processing that originates in primary auditory areas and extends to non-auditory regions, mainly within frontal cortex, in a range of motor, premotor, and prefrontal regions (Davis and Johnsrude, 2007; Hickok and Poeppel, 2007; Rauschecker and Scott, 2009). In non-human primates, based on reports of high level of connection between the auditory and frontal cortex, it has been proposed that frontal regions responsive to auditory material should be considered as part of the auditory system (Hackett et al., 1999; Kaas et al., 1999; Romanski et al., 1999).

FUNCTIONAL SPECIALIZATION OF THE AUDITORY CORTEX RESPONSE: ACOUSTIC COMPLEXITY EFFECTS

As an alternative to the classical division of auditory stimuli into typical categories like pure tones, noise, voices, and music, we explored how acoustic variations along the temporal and spectral dimensions were represented at the cortical level. This approach for defining auditory material is an efficient and comprehensive characterization of sounds that can be considered as a complement to the more typically studied categorical effects. Possibly, certain aspects of human auditory processes might be better characterized in terms of their capacity to analyze acoustic features rather than having differential sensitivity to typical sound categories. In a meta-analysis Rivier and Clarke (1997) found no clear functional specialization in non-primary auditory fields for a range of complex sound categories, showing that processing sounds of different categories such as noise, words, and music, elicited activity in multiple non-primary fields around HG with no emergence of a specific organizational pattern. Similarly, Griffiths and Warren (2002) reported that activity within the PT, an auditory association region, is not spatially organized according to sound categories such as music, speech or, environmental sounds.

By classifying sounds according to their variations in time and frequency, we isolated different levels of auditory complexity, suggesting a within-hemisphere functional segregation with anterior STG and STS more sensitive to changes in the temporal domain and posterior regions (PT and posterolateral STG) more sensitive to changes in along the spectral dimension. Interestingly, a partial overlap was observed between regions sensitive to temporal and spectral changes in lateral HG, suggesting great sensitivity to variations in acoustic properties within this region, consistent with a recent report of strongest sensitivity to stimulus acoustic features within HG (Okada et al., 2010).

Our observation of differential sensitivity to temporal and spectral features can be interpreted in the light of previous findings. First, in the animal literature, a within-hemisphere model of spectral and temporal processing in the auditory cortex has been proposed (Bendor and Wang, 2008). This scheme suggests two streams of processing originating from primary auditory cortex; an anterior pathway sensitive to temporal changes and a lateral pathway responsive to spectral changes. More precise temporal coding is seen as one progresses from primary to anterior auditory regions in primates (Bendor and Wang, 2007) and greater sensitivity to temporal modulations in anterior non-primary auditory fields is also observed in cats (Tian and Rauschecker, 1994). Possibly, a longer integration window in anterior auditory fields could underlie complex temporal processing (Bendor and Wang, 2008). As regards spectral processing, increasing sensitivity to broadband spectrum noise compared to single tones has been observed in lateral and posterior auditory fields in non-human primates (Rauschecker and Tian, 2004; Petkov et al., 2006). Furthermore, given that the neurons within these regions show strong tuning to bandwidth and frequency, some have suggested their involvement in the early stages of spectral analysis of complex sounds (Rauschecker and Tian, 2004). In our study, sensitivity to temporal changes was observed in anterior temporal regions, while, in response to changes along the spectral dimension, we

mainly observed response selectivity in posterolateral auditory fields. Our results therefore seem to be consistent with previous animal studies.

Second, cortical response specificity to spectral and temporal processing has also been studied in humans. Whereas some studies reported no clear functional segregation between responses to spectral and temporal cues (Hall et al., 2002) or observed neuronal populations tuned to specific combinations of spectro-temporal cues (Schönwiesner and Zatorre, 2009), other studies found the sorts of specific sensitivity to spectral vs. temporal features in human auditory cortex we observed in our meta-analysis. For instance, lateral HG and anterolateral PT activity have been reported in association with fine spectral structure analysis (Warren et al., 2005) and change detection of complex harmonic tones involved the posterior STG and lateral PT (Schönwiesner et al., 2007). Additionally, recent studies examining effective connectivity effects among auditory regions reported that spectral envelope analysis follows a serial pathway from HG to PT and then to the STS (Griffiths et al., 2007; Kumar et al., 2007). Conversely for temporal complexity effects, a stream of processing from primary auditory cortex to anterior STG has been observed for auditory pattern analysis such as dynamic pitch variation (Griffiths et al., 1998). Similarly, significant effects of temporal modulation have been reported in anterior non-primary auditory fields (Hall et al., 2000). Some studies therefore report patterns of activity consistent with the current findings, albeit separately for spectral and temporal features.

A more frequently observed feature of spectral vs. temporal processing is between-hemisphere functional specialization. Most studies observed slight but significant lateralization effects with a left-lateralized response to temporal information and right-lateralized activity to spectral information (Zatorre and Belin, 2001; Schönwiesner et al., 2005; Jamison et al., 2006; Obleser et al., 2008). In the current study, lateralization effects were not seen with regard to complexity. However, at higher processing levels, leftward lateralization for speech was observed. Others studies failing to demonstrate the expected lateralization proposed that early stages of processing involve bilateral auditory cortex and that higher cognitive functions, such as speech processing, also rely on these regions but involve more extensive regions in the dominant hemisphere (Langers et al., 2003). Alternatively, Tervaniemi and Hugdahl (2003) reviewed studies showing that response lateralization within the auditory cortex is dependent on sound structure as well as the acoustic background they are presented in. For instance, reduced or absent hemispheric specialization for speech sounds has been reported when the amount of formant structure is not sufficient to establish phoneme categorization (Rinne et al., 1999) or when sounds are presented in noise (Shtyrov et al., 1998). Stimulus heterogeneity among the different experiments included in our meta-analysis could explain why we did not observe asymmetrical hemispheric effects.

To summarize, our meta-analysis demonstrates a clear within-hemisphere functional segregation related to spectral and temporal processing in human auditory cortex, consistent with the known organization of non-human primate auditory system. That such clear spectral vs. temporal complexity gradients are observed (Figure 5), while very few of the included studies have explicitly addressed this issue, illustrates the power of the meta-analysis

approach for human neuroimaging studies. Based on the observed regional functional segregation, we argue that acoustic complexity could well represent a relevant stimulus dimension upon which to identify response segregation within the auditory system. Complexity and categorical effects could therefore be considered as two complementary approaches to more fully characterizing the underlying nature of auditory regional functional specialization.

REFERENCES

- Belin, P. (2006). Voice processing in human and non-human primates. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 361, 2091–2107.
- Belin, P., Zatorre, R. J., and Ahad, P. (2002). Human temporal-lobe response to vocal sounds. *Brain Res. Cogn. Brain Res.* 13, 17–26.
- Belin, P., Zatorre, R. J., Hoge, R., Evans, A. C., and Pike, B. (1999). Event-related fMRI of the auditory cortex. *Neuroimage* 10, 417–429.
- Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., and Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature* 403, 309–312.
- Belin, P., Zilbovicius, M., Crozier, S., Thivard, L., Fontaine, A., Masure, M.-C., and Samson, Y. (1998). Lateralization of speech and auditory temporal processing. *J. Cogn. Neurosci.* 10, 4, 536–540.
- Bendor, D., and Wang, X. (2007). Differential neural coding of acoustic flutter within primate auditory cortex. *Nat. Neurosci.* 10, 763–771.
- Bendor, D., and Wang, X. (2008). Neural response properties of primary, rostral, and rostrotemporal core fields in the auditory cortex of marmoset monkeys. *J. Neurophysiol.* 100, 888–906.
- Benson, R. R., Richardson, M., Whalen, D. H., and Lai, S. (2006). Phonetic processing areas revealed by sinewave speech and acoustically similar non-speech. *Neuroimage* 31, 342–353.
- Benson, R. R., Whalen, D. H., Richardson, M., Swainson, B., Clark, V. P., Lai, S., and Liberman, A. M. (2001). Parametric dissociating speech and nonspeech perception in the brain using fMRI. *Brain Lang.* 78, 364–396.
- Bilecen, D., Scheffler, K., Schmid, N., Tschopp, K., and Seelig, J. (1998). Tonotopic organization of the human auditory cortex as detected by BOLD-fMRI. *Hear. Res.* 126, 19–27.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S. F., Springer, J. A., Kaufman, J. N., and Possing, E. T. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cereb. Cortex* 10, 512–528.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Rao, S. M., and Cox, R. W. (1996). Function of the left planum temporale in auditory and linguistic processing. *Brain* 119(Pt 4), 1239–1247.
- Binder, J. R., Rao, S. M., Hammeke, T. A., Yetkin, F. Z., Jesmanowicz, A., Bandettini, P. A., Wong, E. C., Estkowski, L. D., Goldstein, M. D., and Houghton, V. M. (1994). Functional magnetic resonance imaging of human auditory cortex. *Ann. Neurol.* 35, 662–672.
- Binder, J. R., Swanson, S. J., Hammeke, T. A., and Sabsevitz, D. S. (2008). A comparison of five fMRI protocols for mapping speech comprehension systems. *Epilepsia* 49, 1980–1997.
- Blood, A. J., Zatorre, R. J., Bermudez, P., and Evans, A. C. (1999). Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nat. Neurosci.* 2, 382–387.
- Brown, S., Martinez, M. J., and Parsons, L. M. (2004). Passive music listening spontaneously engages limbic and paralimbic systems. *Neuroreport* 15, 2033–2037.
- Burton, M. W., and Small, S. L. (2006). Functional neuroanatomy of segmenting speech and nonspeech. *Cortex* 42, 644–651.
- Burton, M. W., Small, S. L., and Blumstein, S. E. (2000). The role of segmentation in phonological processing: an fMRI investigation. *J. Cogn. Neurosci.* 12, 679–690.
- Celsis, P., Boulanouar, K., Doyon, B., Ranjeva, J. P., Berry, I., Nespoulous, J. L., and Chollet, F. (1999). Differential fMRI responses in the left posterior superior temporal gyrus and left supramarginal gyrus to habituation and change detection in syllables and tones. *Neuroimage* 9, 135–144.
- Chen, J. L., Penhune, V. B., and Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cereb. Cortex* 18, 2844–2854.
- Chubb, C., and Sperling, G. (1988). Drift-balanced random stimuli: a general basis for studying non-Fourier motion perception. *J. Opt. Soc. Am. A* 5, 1986–2007.
- Dalla Barba, G., Parlato, V., Jobert, A., Samson, Y., and Pappata, S. (1998). Cortical networks implicated in semantic and episodic memory: common or unique? *Cortex* 34, 547–561.
- Davis, M. H., and Johnsrude, I. S. (2003). Hierarchical processing in spoken language comprehension. *J. Neurosci.* 23, 3423–3431.
- Davis, M. H., and Johnsrude, I. S. (2007). Hearing speech sounds: top-down influences on the interface between audition and speech perception. *Hear. Res.* 229, 132–147.
- Dehaene-Lambertz, G., Pallier, C., Serniclaes, W., Sprenger-Charolles, L., Jobert, A., and Dehaene, S. (2005). Neural correlates of switching from auditory to speech perception. *Neuroimage* 24, 21–33.
- Démonet, J.-F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J., Wise, R., Rascol, A., and Frackowiak, R. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain* 115, 1753–1768.
- De Nil, L. F., Beal, D. S., Lafaille, S. J., Kroll, R. M., Crawley, A. P., and Gracco, V. L. (2008). The effects of simulated stuttering and prolonged speech on the neural activation patterns of stuttering and nonstuttering adults. *Brain Lang.* 107, 114–123.
- Dufour, O., Serniclaes, W., Sprenger-Charolles, L., and Démonet, J. F. (2007). Top-down processes during auditory phoneme categorization in dyslexia: a PET study. *Neuroimage* 34, 1692–1707.
- Engelien, A., Tüscher, O., Hermans, W., Isenberg, N., Eidelberg, D., Frith, C., Stern, E., and Silbersweig, D. (2006). Functional neuroanatomy of non-verbal semantic sound processing in humans. *J. Neural Transm.* 113, 599–608.
- Fecteau, S., Armony, J. L., Joannette, Y., and Belin, P. (2004). Is voice processing species-specific in human auditory cortex? An fMRI study. *Neuroimage* 23, 840–848.
- Fernandez-Espejo, D., Junque, C., Vendrell, P., Bernabeu, M., Roig, T., Bargallo, N., and Mercader, J. M. (2008). Cerebral response to speech in vegetative and minimally conscious states after traumatic brain injury. *Brain Inj.* 22, 882–890.
- Fox, P. T., Parsons, L. M., and Lancaster, J. L. (1998). Beyond the single study: function/location metanalysis in cognitive neuroimaging. *Curr. Opin. Neurobiol.* 8, 178–187.
- Gaab, N., Gabrieli, J. D. E., and Glover, G. H. (2007a). Assessing the influence of scanner background noise on auditory processing. I. An fMRI study comparing three experimental designs with varying degrees of scanner noise. *Hum. Brain Mapp.* 28, 703–720.
- Gaab, N., Gabrieli, J. D. E., and Glover, G. H. (2007b). Assessing the influence of scanner background noise on auditory processing. II. An fMRI study comparing auditory processing in the absence and presence of recorded scanner noise using a sparse design. *Hum. Brain Mapp.* 28, 721–732.
- Giraud, A. L., Kell, C., Thierfelder, C., Sterzer, P., Russ, M. O., Preibisch, C., and Kleinschmidt, A. (2004). Contributions of sensory input, auditory search and verbal comprehension to cortical activity during speech processing. *Cereb. Cortex* 14, 247–255.
- Griffiths, T. D., and Warren, J. D. (2002). The planum temporale as a computational hub. *Trends Neurosci.* 25, 348–353.
- Griffiths, T. D., and Warren, J. D. (2004). What is an auditory object? *Nat. Rev. Neurosci.* 5, 887–892.
- Griffiths, T. D., Buchel, C., Frackowiak, R. S., and Patterson, R. D. (1998). Analysis of temporal structure in sound by the human brain. *Nat. Neurosci.* 1, 422–427.
- Griffiths, T. D., Kumar, S., Warren, J. D., Stewart, L., Stephan, K. E., and Friston, K. J. (2007). Approaches to the cortical analysis of auditory objects. *Hear. Res.* 229, 46–53.
- Hackett, T. A. (2008). Anatomical organization of the auditory cortex. *J. Am. Acad. Audiol.* 19, 774–779.
- Hackett, T. A., Preuss, T. M., and Kaas, J. H. (2001). Architectonic identification of the core region in auditory cortex of macaques, chimpanzees, and humans. *J. Comp. Neurol.* 441, 197–222.
- Hackett, T. A., Stepniowska, I., and Kaas, J. H. (1999). Prefrontal connections of the parabelt auditory cortex in macaque monkeys. *Brain Res.* 817, 45–58.
- Hall, D. A., Haggard, M. P., Akeroyd, M. A., Summerfield, A. Q., Palmer, A. R., Elliott, M. R., and Bowtell, R. W. (2000). Modulation and task effects in auditory processing measured using fMRI. *Hum. Brain Mapp.* 10, 107–119.
- Hall, D. A., Johnsrude, I. S., Haggard, M. P., Palmer, A. R., Akeroyd, M. A., and

- Summerfield, A. Q. (2002). Spectral and temporal processing in human auditory cortex. *Cereb. Cortex* 12, 140–149.
- Hart, H. C., Palmer, A. R., and Hall, D. A. (2003). Amplitude and frequency modulated stimuli activate common regions of human auditory cortex. *Cereb. Cortex* 13, 773–781.
- Hart, H. C., Palmer, A. R., and Hall, D. A. (2004). Different areas of human non-primary auditory cortex are activated by sounds with spatial and nonspatial properties. *Hum. Brain Mapp.* 21, 178–190.
- Hertrich, I., Dietrich, S., and Ackermann, H. (2010). Cross-modal interactions during perception of audiovisual speech and nonspeech signals: an fMRI study. *J. Cogn. Neurosci.* 23, 221–237.
- Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402.
- Hugdahl, K., Thomsen, T., Erslund, L., Rimol, L. M., and Niemi, J. (2003). The effects of attention on speech perception: an fMRI study. *Brain Lang.* 85, 37–48.
- Hwang, J. H., Li, C. W., Wu, C. W., Chen, J. H., and Liu, T. C. (2007). Aging effects on the activation of the auditory cortex during binaural speech listening in white noise: an fMRI study. *Audiol. Neurootol.* 12, 285–294.
- 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.
- Jamison, H. L., Watkins, K. E., Bishop, D. V., and Matthews, P. M. (2006). Hemispheric specialization for processing auditory nonspeech stimuli. *Cereb. Cortex* 16, 1266–1275.
- Jäncke, L., Shah, N. J., Posse, S., Grosse-Ryken, M., and Müller-Gärtner, H.-W. (1998). Intensity coding of auditory stimuli: an fMRI study. *Neuropsychologia* 36, 875–883.
- Kaas, J. H., and Hackett, T. A. (2000). Subdivisions of auditory cortex and processing streams in primates. *Proc. Natl. Acad. Sci. U.S.A.* 97, 11793–11799.
- Kaas, J. H., Hackett, T. A., and Tramo, M. J. (1999). Auditory processing in primate cerebral cortex. *Curr. Opin. Neurobiol.* 9, 164–170.
- Koelsch, S., Gunter, T. C., v Cramon, D. Y., Zysset, S., Lohmann, G., and Friederici, A. D. (2002). Bach speaks: a cortical “language-network” serves the processing of music. *Neuroimage* 17, 956–966.
- Kriegstein, K. V., and Giraud, A. L. (2004). Distinct functional substrates along the right superior temporal sulcus for the processing of voices. *Neuroimage* 22, 948–955.
- Kumar, S., Stephan, K. E., Warren, J. D., Friston, K. J., and Griffiths, T. D. (2007). Hierarchical processing of auditory objects in humans. *PLoS Comput. Biol.* 3, e100. doi: 10.1371/journal.pcbi.0030100
- Laird, A. R., Fox, P. M., Price, C. J., Glahn, D. C., Uecker, A. M., Lancaster, J. L., Turkeltaub, P. E., Kochunov, P., and Fox, P. T. (2005). ALE meta-analysis: controlling the false discovery rate and performing statistical contrasts. *Hum. Brain Mapp.* 25, 155–164.
- Lancaster, J. L., Tordesillas-Gutierrez, D., Martinez, M., Salinas, F., Evans, A., Zilles, K., Mazziotta, J. C., and Fox, P. T. (2007). Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. *Hum. Brain Mapp.* 28, 1194–1205.
- Langers, D. R., Backes, W. H., and van Dijk, P. (2003). Spectrotemporal features of the auditory cortex: the activation in response to dynamic ripples. *Neuroimage* 20, 265–275.
- Larsson, J., Landy, M. S., and Heeger, D. J. (2006). Orientation-selective adaptation to first- and second-order patterns in human visual cortex. *J. Neurophysiol.* 95, 862–881.
- Lauter, J. L., Herscovitch, P., Formby, C., and Raichle, M. E. (1985). Tonotopic organization in human auditory cortex revealed by positron emission tomography. *Hear. Res.* 20, 199–205.
- Leaver, A. M., and Rauschecker, J. P. (2010). Cortical representation of natural complex sounds: effects of acoustic features and auditory object category. *J. Neurosci.* 30, 7604–7612.
- Lewis, J. W., Brefczynski, J. A., Phinney, R. E., Janik, J. J., and DeYoe, E. A. (2005). Distinct cortical pathways for processing tool versus animal sounds. *J. Neurosci.* 25, 5148–5158.
- Lewis, J. W., Talkington, W. J., Walker, N. A., Spirou, G. A., Jajosky, A., Frum, C., and Brefczynski-Lewis, J. A. (2009). Human cortical organization for processing vocalizations indicates representation of harmonic structure as a signal attribute. *J. Neurosci.* 29, 2283–2296.
- Liebenthal, E., Binder, J. R., Spitzer, S. M., Possing, E. T., and Medler, D. A. (2005). Neural substrates of phonemic perception. *Cereb. Cortex* 15, 1621–1631.
- Lillywhite, L. M., Saling, M. M., Demutska, A., Masterton, R., Farquharson, S., and Jackson, G. D. (2010). The neural architecture of discourse compression. *Neuropsychologia* 48, 873–879.
- Lockwood, A. H., Salvi, R. J., Coad, M. L., Arnold, S. A., Wack, D. S., Murphy, B. W., and Burkard, R. F. (1999). The functional anatomy of the normal human auditory system: responses to 0.5 and 4.0 kHz tones at varied intensities. *Cereb. Cortex* 9, 65–76.
- 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.
- Müller, R. A., Behen, M. E., Rothermel, R. D., Chugani, D. C., Muzik, O., Mangner, T. J., and Chugani, H. T. (1999). Brain-mapping of language and auditory perception in high-functioning autistic adults: a PET study. *J. Autism Dev. Disord.* 29, 19–31.
- Mummary, C. J., Ashburner, J., Scott, S. K., and Wise, R. J. S. (1999). Functional neuroimaging of speech perception in six normal and two aphasic subjects. *J. Acoust. Soc. Am.* 106, 449–457.
- Mutschler, I., Wieckhorst, B., Speck, O., Schulze-Bonhage, A., Hennig, J., Seifritz, E., and Ball, T. (2010). Time scales of auditory habituation in the amygdala and cerebral cortex. *Cereb. Cortex* 20, 2531–2539.
- Narain, C., Scott, S. K., Wise, R. J., Rosen, S., Leff, A., Iversen, S. D., and Matthews, P. M. (2003). Defining a left-lateralized response specific to intelligible speech using fMRI. *Cereb. Cortex* 13, 1362–1368.
- Obleser, J., Boecker, H., Drzezga, A., Haslinger, B., Hennenlotter, A., Roettinger, M., Eulitz, C., and Rauschecker, J. P. (2006). Vowel sound extraction in anterior superior temporal cortex. *Hum. Brain Mapp.* 27, 562–571.
- Obleser, J., Eisner, F., and Kotz, S. A. (2008). Bilateral speech comprehension reflects differential sensitivity to spectral and temporal features. *J. Neurosci.* 28, 8116–8123.
- Obleser, J., Zimmermann, J., Van Meter, J., and Rauschecker, J. P. (2007). Multiple stages of auditory speech perception reflected in event-related fMRI. *Cereb. Cortex* 17, 2251–2257.
- Okada, K., Rong, F., Venezia, J., Matchin, W., Hsieh, I. H., Saberi, K., Serences, J. T., and Hickok, G. (2010). Hierarchical organization of human auditory cortex: evidence from acoustic invariance in the response to intelligible speech. *Cereb. Cortex* 20, 2486–2495.
- O’Leary, D., Andreasen, N. C., Hurtig, R. R., Hichwa, R. D., Watkins, L., Boles Ponto, L. L., Rogers, M., and Kirchner, P. T. (1996). A positron emission tomography study of binaurally and dichotically presented stimuli: effects of level of language and directed attention. *Brain Lang.* 53, 20–39.
- Patterson, R. D., Uppenkamp, S., Johnsrude, I. S., and Griffiths, T. D. (2002). The processing of temporal pitch and melody information in auditory cortex. *Neuron* 36, 767–776.
- Paulesu, E., Harrison, J., Bron-Cohen, S., Watson, J. D. G., Goldstein, L., Heather, J., Frackowiak, R. S. J., and Frith, C. D. (1995). The physiology of coloured hearing. A PET activation study of colour-word synaesthesia. *Brain* 118, 661–676.
- Peretz, I., Gosselin, N., Belin, P., Zatorre, R. J., Plailly, J., and Tillmann, B. (2009). Music lexical networks: the cortical organization of music recognition. *Ann. N. Y. Acad. Sci.* 1169, 256–265.
- Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., and Raichle, M. E. (1988). Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature* 331, 585–589.
- Petkov, C. I., Kayser, C., Augath, M., and Logothetis, N. K. (2006). Functional imaging reveals numerous fields in the monkey auditory cortex. *PLoS Biol.* 4, e215. doi: 10.1371/journal.pbio.0040215
- Petkov, C. I., Kayser, C., Steudel, T., Whittingstall, K., Augath, M., and Logothetis, N. K. (2008). A voice region in the monkey brain. *Nat. Neurosci.* 11, 367–374.
- Price, C., Wise, R., Ramsay, S., Friston, K., Howard, D., Patterson, K., and Frackowiak R. (1992). Regional response differences within the human auditory cortex when listening to words. *Neurosci. Lett.* 146, 179–182.
- Rauschecker, J. P. (1998). Cortical processing of complex sounds. *Curr. Opin. Neurobiol.* 8, 516–521.
- Rauschecker, J. P., and Scott, S. K. (2009). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat. Neurosci.* 12, 718–724.
- Rauschecker, J. P., and Tian, B. (2000). Mechanisms and streams for processing of “what” and “where” in auditory cortex. *Proc. Natl. Acad. Sci. U.S.A.* 97, 11800–11806.
- Rauschecker, J. P., and Tian, B. (2004). Processing of band-passed noise in the lateral auditory belt cortex of the rhesus monkey. *J. Neurophysiol.* 91, 2578–2589.
- Rauschecker, J. P., Tian, B., and Hauser, M. (1995). Processing of complex sounds in the macaque nonprimary auditory cortex. *Science* 268, 111–114.
- Recanzone, G. H., and Cohen, Y. E. (2010). Serial and parallel processing in the primate auditory cortex revisited. *Behav. Brain Res.* 206, 1–7.
- Reynolds Losin, E. A., Rivera, S. M., O’Hare, E. D., Sowell, E. R., and Pinter, J. D. (2009). Abnormal fMRI activation pattern during story listening in individuals with Down syndrome. *Am. J. Intellect. Dev. Disabil.* 114, 369–380.
- Rimol, L. M., Specht, K., Weis, S., Savoy, R., and Hugdahl, K. (2005). Processing of sub-syllabic speech units in the

- posterior temporal lobe: an fMRI study. *Neuroimage* 26, 1059–1067.
- Rinne, T., Alho, K., Alku, P., Holi, M., Sinkkonen, J., Virtanen, J., Bertrand, O., and Näätänen, R. (1999). Analysis of speech sounds is left-hemisphere predominant at 100–150 ms after sound onset. *Neuroreport* 10, 1113–1117.
- Rivier, F., and Clarke, S. (1997). Cytochrome oxidase, acetylcholinesterase, and NADPH-diaphorase staining in human supratemporal and insular cortex: evidence for multiple auditory areas. *Neuroimage* 6, 288–304.
- Romanski, L. M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P. S., and Rauschecker, J. P. (1999). Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nat. Neurosci.* 2, 1131–1136.
- Sabri, M., Binder, J. R., Desai, R., Medler, D. A., Leitl, M. D., and Liebenthal, E. (2008). Attentional and linguistic interactions in speech perception. *Neuroimage* 39, 1444–1456.
- Schönwiesner, M., Novitski, N., Pakarinen, S., Carlson, S., Tervaniemi, M., and Naatanen, R. (2007). Heschl's gyrus, posterior superior temporal gyrus, and mid-ventrolateral prefrontal cortex have different roles in the detection of acoustic changes. *J. Neurophysiol.* 97, 2075–2082.
- Schönwiesner, M., Rubsamen, 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.
- Schönwiesner, M., and Zatorre, R. J. (2009). Spectro-temporal modulation transfer function of single voxels in the human auditory cortex measured with high-resolution fMRI. *Proc. Natl. Acad. Sci. U.S.A.* 106, 14611–14616.
- Schwarzbauer, C., Davis, M. H., Rodd, J. M., and Johnsrude, I. (2006). Interleaved silent steady state (ISSS) imaging: a new sparse imaging method applied to auditory fMRI. *Neuroimage* 29, 774–782.
- Scott, S. K., Blank, C. C., Rosen, S., and Wise, R. J. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain* 123(Pt 12), 2400–2406.
- Scott, S. K., and Johnsrude, I. S. (2003). The neuroanatomical and functional organization of speech perception. *Trends Neurosci.* 26, 100–107.
- Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science* 270, 303–304.
- Shtyrov, Y., Kujala, T., Ahveninen, J., Tervaniemi, M., Alku, P., Ilmoniemi, R. J., and Näätänen, R. (1998). Background acoustic noise and the hemispheric lateralization of speech processing in the human brain: magnetic mismatch negativity study. *Neurosci. Lett.* 251, 141–144.
- Specht, K., and Reul, J. (2003). Functional segregation of the temporal lobes into highly differentiated subsystems for auditory perception: an auditory rapid event-related fMRI-task. *Neuroimage* 20, 1944–1954.
- Steinbrink, C., Ackermann, H., Lachmann, T., and Riecker, A. (2009). Contribution of the anterior insula to temporal auditory processing deficits in developmental dyslexia. *Hum. Brain Mapp.* 30, 2401–2411.
- Stevens, A. A. (2004). Dissociating the cortical basis of memory for voices, words and tones. *Brain Res. Cogn. Brain Res.* 18, 162–171.
- Sweet, R. A., Dorph-Petersen, K. A., and Lewis, D. A. (2005). Mapping auditory core, lateral belt, and parabelt cortices in the human superior temporal gyrus. *J. Comp. Neurol.* 491, 270–289.
- Tervaniemi, M., and Hugdahl, K. (2003). Lateralization of auditory-cortex functions. *Brain Res. Brain Res. Rev.* 43, 231–246.
- Thierry, G., Giraud, A. L., and Price, C. (2003). Hemispheric dissociation in access to the human semantic system. *Neuron* 38, 499–506.
- Thivard, L., Belin, P., Zilbovicius, M., Poline, J. B., and Samson Y. (2000). A cortical region sensitive to auditory spectral motion. *Neuroreport* 11, 2969–2972.
- Tian, B., and Rauschecker, J. P. (1994). Processing of frequency-modulated sounds in the cat's anterior auditory field. *J. Neurophysiol.* 71, 1959–1975.
- Tramo, M. J., Shah, G. D., and Braid, L. D. (2002). Functional role of auditory cortex in frequency processing and pitch perception. *J. Neurophysiol.* 87, 122–139.
- Turkeltaub, P. E., Eden, G. F., Jones, K. M., and Zeffiro, T. A. (2002). Meta-analysis of the functional neuroanatomy of single-word reading: method and validation. *Neuroimage* 16(Pt 1), 765–780.
- Uppenkamp, S., Johnsrude, I. S., Norris, D., Marslen-Wilson, W., and Patterson, R. D. (2006). Locating the initial stages of speech-sound processing in human temporal cortex. *Neuroimage* 31, 1284–1296.
- Vouloumanos, A., Kiehl, K. A., Werker, J. F., and Liddle, P. F. (2001). Detection of sounds in the auditory stream: event-related fMRI evidence for differential activation to speech and nonspeech. *J. Cogn. Neurosci.* 13, 994–1005.
- Wager, T. D., Lindquist, M. A., Nichols, T. E., Kober, H., and Van Snellenberg, J. X. (2009). Evaluating the consistency and specificity of neuroimaging data using meta-analysis. *Neuroimage* 45(Suppl.1), S210–S221.
- Wallace, M. N., Johnston, P. W., and Palmer, A. R. (2002). Histochemical identification of cortical areas in the auditory region of the human brain. *Exp. Brain Res.* 143, 499–508.
- Warren, J. D., and Griffiths, T. D. (2003). Distinct mechanisms for processing spatial sequences and pitch sequences in the human auditory brain. *J. Neurosci.* 23, 5799–5804.
- Warren, J. D., Jennings, A. R., and Griffiths, T. D. (2005). Analysis of the spectral envelope of sounds by the human brain. *Neuroimage* 24, 1052–1057.
- Warrier, C., Wong, P., Penhune, V., Zatorre, R., Parrish, T., Abrams, D., and Kraus, N. (2009). Relating structure to function: Heschl's gyrus and acoustic processing. *J. Neurosci.* 29, 61–69.
- Warrier, C. M., and Zatorre, R. J. (2002). Influence of tonal context and timbral variation on perception of pitch. *Percept. Psychophys.* 64, 198–207.
- Warrier, C. M., and Zatorre, R. J. (2004). Right temporal cortex is critical for utilization of melodic contextual cues in a pitch constancy task. *Brain* 127(Pt 7), 1616–1625.
- Wessinger, C. M., VanMeter, J., Tian, B., Van Lare, J., Pekar, J., and Rauschecker, J. P. (2001). Hierarchical organization of the human auditory cortex revealed by functional magnetic resonance imaging. *J. Cogn. Neurosci.* 13, 1–7.
- Wise, R., Chollet, F., Hadar, U., Friston, K., Hoffner, E., and Frackowiak, R. (1991). Distribution of cortical networks involved in word comprehension and word retrieval. *Brain* 114, 1803–1817.
- Woods, D. L., and Alain, C. (2009). Functional imaging of human auditory cortex. *Curr. Opin. Otolaryngol. Head Neck Surg.* 17, 407–411.
- Zaehle, T., Schimidt, C. F., Meyer, M., Baumann, S., Baltes, C., Boesiger, P., and Jancke, L. (2007). Comparison of “silent” clustered and sparse temporal fMRI acquisitions in tonal and speech perception tasks. *Neuroimage* 37, 1195–1204.
- Zatorre, R. J., and Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cereb. Cortex* 11, 946–953.
- 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.
- Zatorre, R. J., Bouffard, M., Ahad, P., and Belin, P. (2002). Where is “where” in the human auditory cortex? *Nat. Neurosci.* 5, 905–909.
- Zatorre, R. J., Evans, A. C., and Meyer, E. (1994). Neurochemical mechanisms underlying melodic perception and memory for pitch. *J. Neurosci.* 14, 1908–1919.
- Zatorre, R. J., Evans, A. C., Meyer, A., and Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science* 256, 846–849.
- Zatorre, R. J., Meyer, E., Gjedde, A., and Evans, A. C. (1996). PET studies of phonetic processing of speech: review, replication and reanalysis. *Cereb. Cortex* 6, 21–30.
- Zevin, J., and McCandliss, B. (2005). Dishabituation of the BOLD response to speech sounds. *Behav. Brain Funct.* 1, 4.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 16 September 2010; accepted: 23 December 2010; published online: 17 January 2011.

Citation: Samson F, Zeffiro TA, Toussaint A and Belin P (2011) Stimulus complexity and categorical effects in human auditory cortex: an Activation Likelihood Estimation meta-analysis. *Front. Psychology* 1:241. doi: 10.3389/fpsyg.2010.00241

This article was submitted to *Frontiers in Auditory Cognitive Neuroscience*, a specialty of *Frontiers in Psychology*.

Copyright © 2011 Samson, Zeffiro, Toussaint and Belin. This is an open-access article subject to an exclusive license agreement between the authors and the Frontiers Research Foundation, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited.

APPENDIX

Table A1 | Selection of foci for the subsamples: Vocal sounds (20 contrasts, 165 foci) Intelligible speech (27 contrasts, 166 foci), Complex non-speech (13 contrasts, 76 foci) and Multiple frequency, irregular change (10 contrasts, 70 foci).

Reference	Table/ figure	Category	Frequency composition	Changes over time	Stimuli of interest	Baseline	Foci	Vocal	Intelligible speech	Complex non-speech	Multiple, irregular change
Belin et al. (1998)	Table 1, p. 537	–	Multiple	Irregular	Rapid formant transition	Silence	3			X	X
		–	Multiple	Irregular	Extended formant transition	Silence	2			X	
Belin et al. (1999)	Text, p. 422–423	–	Multiple	Unchanging	Harmonic complex sound	Silence	4				
Benson et al. (2001)	Table 1, p. 372	Music	–	–	Note/chord/chord progression	Rest	27				
	Table 2, p. 373	Vocal	Multiple	Irregular	Vowel/consonant– vowels/syllables	Rest	24	X			
Binder et al. (1996)	Table 1, p. 1244	Vocal	Multiple	Irregular	Words	Tones sequence	1		X		
Binder et al. (2000)	Table 2, p. 518	–	Multiple	Regular	Sequence of tones	Noise	11				
		Vocal	Multiple	Irregular	Words	Sequence of tones	8		X		
		Vocal	Multiple	Irregular	Pseudowords	Sequence of tones	8	X		X	
Binder et al. (2008)	Appendix	Vocal	Multiple	Irregular	Words	Sequence of tones	7			X	X
		Vocal	Multiple	Irregular	Words	Rest	12		X		
		Vocal	Multiple	Irregular	Words	Tones	2				
Blood et al. (1999)	Table 2, p. 384	Music	–	–	Melodies	Noise	3				
		Music	–	–	Melodies	Noise	3				
Brown et al. (2004)	Table 1, p. 2035	Music	Multiple	Irregular	Music	Rest	21				
Burton et al. (2000)	Table 2, p. 682	Vocal	Multiple	Irregular	Words	Tone	3		X		
		Vocal	Multiple	Irregular	Words	Tone	6	X			
Burton and Small (2006)	Table 2, p. 647	Vocal	Multiple	Irregular	Words	Tone	4		X		
Celsis et al. (1999)	Table 1, p. 138	Tone	Single	Unchanging	Tone	Rest	1				
		Tone	Single	Unchanging	Tone	Rest	1				
		Vocal	Multiple	Irregular	Syllable	Rest	3	X			
		Tone	Single	Unchanging	Tone	Rest	3				
		–	Multiple	Unchanging	Square wave tone	Rest	3				
Chen et al. (2008)	Table 2, p. 2849	Vocal	Multiple	Irregular	Syllable	Rest	3				
		Music	Multiple	Irregular	Musical rhythms	Silence	9				X
Dalla Barba et al. (1998)	Table 1, p. 552	Vocal	Multiple	Irregular	Words	Rest	5		X		

(Continued)

Table A1 | Continued

Reference	Table/figure	Category	Frequency composition	Changes over time	Stimuli of interest	Baseline	Foci	Vocal	Intelligible speech	Complex non-speech	Multiple, irregular change
De Nil et al. (2008)	Table 2, p. 119	Vocal	Multiple	Irregular	Words	Silence	1		X		
Démonet et al. (1992)	Table 4, p. 1758	Vocal	Multiple	Irregular	Phonemes (three or four syllables non-words)	Tones	7	X			
Engelien et al. (2006)	Table 1, p. 603	Vocal	Multiple	Irregular	Words	Tones	14		X		
		–	Multiple	Irregular	Meaningful non-verbal sounds	Rest	8			X	X
Engelien et al. (2006)	Table 1, p. 603	–	Multiple	Irregular	Meaningless non-verbal sounds	Rest	14			X	
		Vocal	Multiple	Irregular	Narrative sentences	Silence	2		X		
Fernandez-Espejo et al. (2008)	Table 2, p. 887	Vocal	Multiple	Irregular	Narrative sentences	Silence	2		X		
Gaab et al. (2007a)	Table 1, p. 710	Vocal	Multiple	Irregular	Words	Silence	7		X		X
Gaab et al. (2007b)	Table 1, p. 727	Vocal	Multiple	Irregular	Words	Silence	9	X			
		Vocal	Multiple	Irregular	Words	Silence	5	X			
Gaub et al. (2007b)	Table 1, p. 727	–	Multiple	Irregular	Recorded scanner noise	Silence	2			X	
		Vocal	–	–	Speech (sentences)	Temporally matched complex noise	5		X		
Giraud et al. (2004)	Table 1, p. 250	Vocal	–	–	Speech (sentences)	Temporally matched complex noise	5		X		
Griffiths et al. (1998)	Table 1, p. 424	Music	–	–	Melodies	Sequence of tones	4				
Hall et al. (2000)	Table 1, p. 114	–	Single	Regular	Amplitude and frequency modulated tone	Static tone	2				
Hall et al. (2002)	Table 2, p. 144	–	Multiple	Unchanging	Harmonic complex tone	Single tone	5				
		–	Single	Regular	Frequency modulated tone	Static tone	10				
Hart et al. (2003)	Text, p. 778	–	Single	Regular	Amplitude and frequency modulated tone	Unmodulated tone	2				
Hart et al. (2004)	Table 1, p. 182	Tone	Single	Unchanging	Tone	Silence	4				
		–	Single	Regular	Frequency modulated tone	Silence	6				
Hertrich et al. (2010)	Table S1	Vocal	Multiple	Irregular	Syllables	Silence	10				
		–	Single	Regular	Single formants sweep	Silence	8				
Hugdahl et al. (2003)	Table 3, p. 43	Vocal	Multiple	Irregular	Vowel	Silence	4				

(Continued)

Table A1 | Continued

Reference	Table/ figure	Category	Frequency composition	Changes over time	Stimuli of interest	Baseline	Foci	Vocal	Intelligible speech	Complex non-speech Multiple, irregular change	
Hwang et al. (2007)	Table 4, p. 43	Vocal	Multiple	Irregular	Pseudowords	Silence	3	X		X	
	Table 5, p. 43	Vocal	Multiple	Irregular	Words	Silence	2		X		
	Table 2, p. 289	Vocal	Multiple	Irregular	Sentences	Quiet	7		X	X	
	Jäncke et al. (1998)	Table 1, p. 878	Vocal	Multiple	Irregular	Syllables	Silence	9	X		
		Table 2, p. 878	Vocal Vocal –	Multiple Multiple Multiple	Irregular Irregular Regular	Syllables Syllables Tone sequence	Silence Silence Silence	9 9 4			
Lillywhite et al. (2010)	Table 1, p. 875	–	Multiple	Regular	Tone sequence	Silence	6				
	Table 3, p. 24	–	Multiple	Regular	Tone sequence	Silence	6				
Müller et al. (1999)	Table 1, p. 452	Vocal	Multiple	Irregular	Story listening	White noise	2	X	X		
	Table 3, p. 24	–	Multiple	Regular	Sequence of tones	Rest	6				
Mummery et al. (1999)	Table 1, p. 452	Vocal Vocal	–	–	Sentences Speech (words)	Rest Signal	5 7	X	X X		
	Supp. Material	Music	Multiple	Irregular	Melodies	Rest Rest	94				
O'Leary et al. (1996)	Table 2, p. 27	Vocal	Multiple	Irregular	Words	Tone	2		X		
	–	–	Multiple	Irregular	Environmental sounds	Tone	2			X	
Obleser et al. (2006)	Table 2, p. 566	Vocal	Multiple	Irregular	Vowel	Band-pass noise	4	X			
Paulesu et al. (1995)	Table 2, p. 667	Vocal	Multiple	Irregular	Words	Tones	5		X		
Peretz et al. (2009)	Table 1, p. 263	Music	–	–	Melodies	Random tone sequence	8				
Petersen et al. (1988)	Table 2, p. 585	Vocal	Multiple	Irregular	Words	Silence (fixation point)	6		X		
Price et al. (1992)	Figure 1, p. 180	Vocal	Multiple	Irregular	Nouns	Rest	6	X	X		
Reynolds Losin et al. (2009)	Table 1, p. 374	Vocal	Multiple	Irregular	Story listening	Rest	3		X		
Rimol et al. (2005)	Table 1, p. 1063	Vocal	Multiple	Irregular	Consonant	Rest	8	X			
	–	Vocal	Multiple	Irregular	Syllables	Rest	6				
	–	Noise	Multiple	Unchanging	Noise (white, brown, pink)	Rest	12				
	–	Noise	Multiple	Unchanging	Noise (white, brown, pink)	Rest	8				
	Table 2, p. 1064	Vocal	Multiple	Irregular	Consonant	Noise	1				
–	Vocal	Multiple	Irregular	Syllable	Noise	3					

(Continued)

Table A1 | Continued

Reference	Table/figure	Category	Frequency composition	Changes over time	Stimuli of interest	Baseline	Foci	Vocal	Intelligible speech	Complex non-speech	Multiple, irregular change
Schwarzbauer et al. (2006)	Table 1, p. 779	Vocal	–	–	Sentences	Signal correlated noise	12	X	X		
		Vocal	–	–	Sentences	Signal correlated noise	6				
Schönwiesner et al. (2005)	Text, p. 1523–1525	–	Single	Regular	Parametric increase in temporal rate change	Silence	2				
		–	Multiple	Unchanging	Parametric increase in spectral components	Silence	2				
Specht and Reul (2003)	Table 1, p. 1950	Tone	Single	Unchanging	Tones	Rest	5				
		–	Multiple	Irregular	Sounds of animal and instruments	Rest	9			X	X
Steinbrink et al. (2009)	Table 2, p. 1951	Vocal	Multiple	Irregular	Words	Rest	18	X	X		
		–	Multiple	Irregular	Sounds of animal and instruments	Tones	7			X	
Steinbrink et al. (2009)	Table 1, p. 2406	Vocal	Multiple	Irregular	Increase rate of syllable presentation	Silence	12				X
		Vocal	Multiple	Irregular	Increase rate of click sequence	Silence	7				
Stevens (2004)	Table 1, p. 166	–	Multiple	Regular	Voice	Silence	8				
		Vocal	Multiple	Irregular	Words	Tone	13		X		
Thivard et al. (2000)	Table 1, p. 2971	Vocal	Multiple	Irregular	Words	Tone	5				
		–	Multiple	Unchanging	Static complex sound	Silence	6				
		–	Multiple	Irregular	Complex sound modulated (motion)	Silence	6			X	
Vouloumanos et al. (2001)	Table 1, p. 998	–	Single	Irregular	Complex sound modulated (motion)	Static complex sound	4				
		Vocal	Multiple	Irregular	Words	Tones	8		X		
Warren and Griffiths (2003)	Table 1, p. 5802	–	Multiple	Irregular	Non-speech complex sinewave sounds	Tones	5			X	X
		–	Multiple	Irregular	Changing pitch sequences	Fixed pitch sequences	4				
Warrier et al. (2009)	Table 2, p. 65	–	Single	Regular	Parametric increase in temporal rate change	Silence	6				

(Continued)

Table A1 | Continued

Reference	Table/ figure	Category	Frequency composition	Changes over time	Stimuli of interest	Baseline	Foci	Vocal	Intelligible speech	Complex non-speech Multiple, irregular change
		–	Multiple	Unchanging	Parametric increase in spectral components	Silence	6			
Wise et al. (1991)	Table 2, p. 1808	Vocal	Multiple	Irregular	Nouns	Rest	6	X	X	
Zaehle et al. (2007)	Table 1, p. 1201	Vocal	Multiple	Irregular	Sentences	Silence	2		X	
		Vocal	Multiple	Irregular	Sentences	Silence	2			
		Tone	Single	Unchanging	Tones	Silence	2			
		Tone	Single	Unchanging	Tones	Silence	2			
		Vocal	Multiple	Irregular	Sentences	Tones	2			
		Vocal	Multiple	Irregular	Sentences	Tones	2			
	Table 2, p. 1201	Vocal	Multiple	Irregular	Sentences	Silence	2			
		Vocal	Multiple	Irregular	Sentences	Silence	2			
		Tone	Single	Unchanging	Tones	Silence	2			
		Tone	Single	Unchanging	Tones	Silence	2			
		Vocal	Multiple	Irregular	Sentences	Tones	2			
		Vocal	Multiple	Irregular	Sentences	Tones	2			
Zatorre et al. (1992)	Table 2, p. 847	Noise	Multiple	Unchanging	White noise bursts	Silence	8			
Zatorre et al. (1994)	Table 2, p. 1911	Music	–	–	Syllables Melodies	Noise Noise	6 2	X		
Zatorre et al. (1996)	Table 6, p. 26	Vocal	Multiple	Irregular	Syllables	Noise	6	X		
Zatorre et al. (2002)	Table 1, p. 907	Noise	Multiple	Irregular	Noise (reversed environmental sounds)	Silence	3			
Zatorre and Belin (2001)	Table 1, p. 948	–	Single	Regular	Tone sequence with increasing rate of presentation	Silence	2			
		–	Multiple	Unchanging	Tone sequence with increasing number of spectral elements	Silence	3			
Zevin and McCandliss (2005)	Table 1	Vocal	Multiple	Irregular	Syllables	Silence	8	X		X