ARTICLE IN PRESS

Hearing Research xxx (2013) 1–13

Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Research paper

Stimulus-dependent activations and attention-related modulations in the auditory cortex: A meta-analysis of fMRI studies

Kimmo Alho [a,b,*], Teemu Rinne [b], Timothy J. Herron [c], David L. Woods [c,d]

[a] *Helsinki Collegium for Advanced Studies, University of Helsinki, PO Box 4, FI 00014 Helsinki, Finland*
[b] *Institute of Behavioural Sciences, University of Helsinki, PO Box 9, FI 00014 Helsinki, Finland*
[c] *Veterans Affairs Northern California Health Care System (VANCHSC), Research Building R4 VAMC, 150 Muir Road, Martinez, CA 94553-4612, USA*
[d] *Department of Neurology, Center for Neuroscience, and Center for Mind and Brain, University of California at Davis, 267 Cousteau Place, Davis, CA 95618, USA*

ARTICLE INFO

Article history:
Received 17 May 2013
Received in revised form
22 July 2013
Accepted 1 August 2013
Available online xxx

ABSTRACT

We meta-analyzed 115 functional magnetic resonance imaging (fMRI) studies reporting auditory-cortex (AC) coordinates for activations related to active and passive processing of pitch and spatial location of non-speech sounds, as well as to the active and passive speech and voice processing. We aimed at revealing any systematic differences between AC surface locations of these activations by statistically analyzing the activation loci using the open-source Matlab toolbox VAMCA (Visualization and Meta-analysis on Cortical Anatomy). AC activations associated with pitch processing (e.g., active or passive listening to tones with a varying vs. fixed pitch) had median loci in the middle superior temporal gyrus (STG), lateral to Heschl's gyrus. However, median loci of activations due to the processing of infrequent pitch changes in a tone stream were centered in the STG or planum temporale (PT), significantly posterior to the median loci for other types of pitch processing. Median loci of attention-related modulations due to focused attention to pitch (e.g., attending selectively to low or high tones delivered in concurrent sequences) were, in turn, centered in the STG or superior temporal sulcus (STS), posterior to median loci for passive pitch processing. Activations due to spatial processing were centered in the posterior STG or PT, significantly posterior to pitch processing loci (processing of infrequent pitch changes excluded). In the right-hemisphere AC, the median locus of spatial attention-related modulations was in the STS, significantly inferior to the median locus for passive spatial processing. Activations associated with speech processing and those associated with voice processing had indistinguishable median loci at the border of mid-STG and mid-STS. Median loci of attention-related modulations due to attention to speech were in the same mid-STG/STS region. Thus, while attention to the pitch or location of non-speech sounds seems to recruit AC areas less involved in passive pitch or location processing, focused attention to speech predominantly enhances activations in regions that already respond to human vocalizations during passive listening. This suggests that distinct attention mechanisms might be engaged by attention to speech and attention to more elemental auditory features such as tone pitch or location.

This article is part of a Special Issue entitled <Human Auditory Neuroimaging>.

1. Introduction

Measurement of blood oxygenation level dependent (BOLD) brain responses with functional magnetic resonance imaging (fMRI) allows millimeter-scale localization of human brain activity. Hundreds of fMRI studies have investigated different aspects of auditory processing and meta-analyses have been performed comparing spatial and nonspatial auditory processing (Arnott et al., 2004) and comparing the processing of speech sounds, human vocalizations and other complex sounds (Frühholz and Granjean, 2013; Samson et al., 2011; Schirmer et al, 2012; Turkeltaub and

Abbreviations: 2D, two-dimensional; 3D, three-dimensional; AC, auditory cortex; AG, angular gyrus; ALE, activation likelihood estimate; BOLD, blood oxygenation level dependent; ERP, event-related brain potential; FM, frequency-modulated; fMRI, functional magnetic resonance imaging; HG, Heschl's gyrus; MEG, magnetoencephalography; MMN, mismatch negativity; MNI, Montreal Neurological Institute; MTG, middle temporal gyrus; PET, positron emission tomography; PP, planum polare; PT, planum temporale; SMG, supramarginal gyrus; STG, superior temporal gyrus; STS, superior temporal sulcus; VAMCA, Visualization and Meta-analysis on Cortical Anatomy

* Corresponding author. Helsinki Collegium for Advanced Studies, University of Helsinki, PO Box 4, FI 00014 Helsinki, Finland. Tel.: +358 405577578 (mobile).

E-mail addresses: kimmo.alho@helsinki.fi (K. Alho), teemu.rinne@helsinki.fi (T. Rinne), tjherron@ebire.org (T.J. Herron), dlwoods@ucdavis.edu (D.L. Woods).

0378-5955/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.heares.2013.08.001

Coslett, 2010; Vigneau et al., 2006). Here, we performed a comprehensive meta-analysis of cortical surface locations of auditory-cortex (AC) activations reported in 115 fMRI studies investigating the passive or attentive processing of pitch or location of tones and other non-speech sounds and the passive or attentive processing of speech sounds and other human vocalizations.

The aim of the present meta-analysis was to reveal consistencies in the locations of AC activations across these numerous auditory fMRI studies. A central issue was whether the processing of different auditory features (i.e., pitch vs. location) and different types of sounds (i.e., speech vs. non-speech sounds) occurred systematically in different regions of AC. We also compared loci of activations for pitch, spatial, and speech or voice processing during passive listening with locations of attention-related modulations associated with attention to pitch or location of tones or with attention to speech to evaluate whether attention merely enhanced activations observed in passive conditions or whether additional regions were recruited when sounds were attended.

We meta-analyzed the AC activation loci using VAMCA (Visualization and Meta-analysis on Cortical Anatomy for details, see Methods), which quantifies differences in median cortical surface activation loci between two experimental conditions across different studies. Unlike standard methods, VAMCA uses no smoothing of the reported stereotaxic coordinates for activation loci. Rather it performs permutation testing on group membership of the anatomical coordinates of activation loci to establish the significance of differences in median activation loci between different conditions. Thus, VAMCA avoids both the non-trivial statistical problems of ensuring that density levels are significant (Costafreda et al., 2009) and the issue of selecting appropriate smoothing kernel size (Eickhoff et al., 2009). Finally, and importantly in our view, VAMCA performs meta-analyses in normalized 2D cortical surface space. This approach enables the results to be visualized relative to the surface anatomy of auditory cortex (Kang et al., 2012) and evaluates the statistical significance of differences in median loci using distance on the cortical surface as distinct from differences in volumetric space (Anticevic et al., 2008; Kang et al., 2007).

Pitch processing in the human brain has been investigated in numerous fMRI studies. These studies have compared, in active or passive listening conditions, activations elicited by a sequence of tones or noise bursts of varying pitch with activations elicited by a sequence of tones or noise bursts with a fixed pitch (e.g., Barrett and Hall, 2006; Hyde, 2008; Opitz et al., 2005; Patterson et al., 2002; Warren and Griffiths, 2003), activations to frequency-modulated (FM) tones with activations to non-FM tones (e.g., Hall et al., 2002; Hart et al., 2004), and activations to tones or noise bursts that produce a salient perception of pitch with activations to sounds that produce weak or absent pitch (e.g., Barker et al., 2011; Binder et al., 2000; Hall et al., 2005; Penagos et al., 2004). Other studies have compared activations during active pitch processing, for example, pitch discrimination tasks or pitch working-memory tasks with activations during other auditory tasks (e.g., Alain et al., 2001; Brechmann and Scheich, 2005; Müller et al., 2001).

Likewise, fMRI studies of spatial auditory processing in active or passive listening conditions have compared activations to moving sounds or to sounds with varying azimuthal locations with activations to stationary sounds (e.g., Barrett and Hall, 2006; Brunetti et al., 2005; Griffiths et al., 2000; Hart et al., 2004; Krumbholz et al., 2005; Pavani et al., 2002; Warren and Griffiths, 2003; Warren et al., 2002), activations to sounds from a well-defined location with activations to sounds with a diffuse source (e.g., Budd et al., 2003; Hall et al., 2005), and analyzed activations during spatial discrimination or spatial working-memory tasks (e.g., Bidet-Caulet et al., 2005; Martinkauppi et al., 2000). A previous meta-analysis of such studies suggested that AC activations associated with spatial auditory processing occur predominantly in posterior AC regions, whereas activations for non-spatial auditory processing are distributed more evenly along the superior temporal cortex (Arnott et al., 2004).

Another line of auditory fMRI research, motivated by event-related brain potential (ERP) studies, has focused on processing of infrequent pitch changes occurring in a to-be-ignored sequence of tones during attention to other sensory inputs (e.g., Doeller et al., 2003; Opitz et al., 1999, 2002; Sabri et al., 2004; Sevostianov et al., 2002). Other fMRI experiments have used an active form of the oddball paradigm, i.e., where deviant tones among standard-pitch tones are to be detected (e.g., Kiehl et al., 2001, 2005; Müller et al., 2003; Sevostianov et al., 2002; Stevens et al., 2000; Yoshiura et al., 1999). In ERPs, such target tones elicit the P3b (P300) response associated with auditory discrimination and context updating (for a review, see Näätänen et al., 2007). However, ERP and magnetoencephalography (MEG) studies also show that even changes in unattended repeating sounds elicit a mismatch negativity (MMN) response in AC. The MMN is apparently elicited by a mismatch between the sensory-memory representation of repeated auditory attributes, for example, the pitch and location of a frequent “standard” tone, and an infrequent auditory event, for example, a “deviant” tone with a different pitch or location. Such irregular events may also initiate involuntary attention, associated with the P3a response which often follows the MMN (for reviews, see Escera et al., 2000; Näätänen et al., 2007). The P3a response is also partly generated in AC (Alho et al., 1998). Therefore AC responses measured with fMRI to deviant sounds are also likely to include significant activity related to the involuntary attention elicited by these sounds even when participants are not required to attend to the auditory input (Escera et al., 2000).

Numerous fMRI studies have also focused on speech processing. Many of these studies have compared AC activations to speech sounds (e.g., vowels, consonants, syllables, words, pseudowords, and sine-wave speech) with activations to non-speech sounds (e.g., noise, tones, and chords) during passive or active listening (e.g., Benson et al., 2001, 2006; Binder et al., 2000; Burton et al., 2000; Burton and Small, 2006; Dick et al., 2007; Jäncke et al., 2002a, 2002b; Liebenthal et al., 2005; LoCasto et al., 2004; Obleser et al., 2006; Rimol et al., 2005, 2006; Specht and Reul, 2003; Uppenkamp et al., 2006; Zaehle et al., 2004). In addition, studies have contrasted AC activations to intelligible speech with activations to unintelligible degraded speech or speech-enveloped noise (Davis and Johnsrude, 2003; Giraud et al., 2004; Narain et al., 2004; Obleser et al., 2007, 2008; Specht et al., 2005). Some studies have contrasted AC activations to native speech with activations produced by incomprehensible foreign languages or non-native phonetic contrasts (Jacquemot et al., 2003; Schlosser et al, 1998), or studied activations associated with the detection of speech sounds (Alain et al., 2005; Binder et al., 2004). Related studies have analyzed AC activations during working memory for pseudowords (Strand et al., 2008) and when recognizing spoken sentences (von Kriegstein et al., 2003). Previous meta-analyses of brain imaging studies on speech processing typically show activations associated with phonological and semantic processing along the superior temporal gyrus (STG), predominantly in the left hemisphere (Turkeltaub and Coslett, 2010; Vigneau et al., 2006).

Parallel fMRI studies have investigated the processing of the human voice in active or passive listening conditions. These studies have compared AC activations to vocal sounds (speech and non-speech) with activations to non-vocal environmental sounds, animal cries, white noise, and scrambled vocalizations in passive listening conditions (e.g., Belin et al., 2000; Bélizaire et al., 2007; Fecteau et al., 2004). Other studies of voice processing have