

Published in final edited form as:

*Hear Res.* 2014 January ; 307: . doi:10.1016/j.heares.2013.06.010.

## Using neuroimaging to understand the cortical mechanisms of auditory selective attention

Adrian KC Lee<sup>1,2</sup>, Eric Larson<sup>1</sup>, Ross K Maddox<sup>1</sup>, and Barbara G Shinn-Cunningham<sup>3,4</sup>

<sup>1</sup>Institute for Learning and Brain Sciences, University of Washington

<sup>2</sup>Department of Speech & Hearing Sciences, University of Washington

<sup>3</sup>Center for Computational Neuroscience and Neural Technology, Boston University, Boston, MA 02215, USA

<sup>4</sup>Department of Biomedical Engineering, Boston University, Boston, MA 02215, USA

### Abstract

Over the last four decades, a range of different neuroimaging tools have been used to study human auditory attention, spanning from classic event-related potential studies using electroencephalography to modern multimodal imaging approaches (e.g., combining anatomical information based on magnetic resonance imaging with magneto- and electroencephalography). This review begins by exploring the different strengths and limitations inherent to different neuroimaging methods, and then outlines some common behavioral paradigms that have been adopted to study auditory attention. We argue that in order to design a neuroimaging experiment that produces interpretable, unambiguous results, the experimenter must not only have a deep appreciation of the imaging technique employed, but also a sophisticated understanding of perception and behavior. Only with the proper caveats in mind can one begin to infer how the cortex supports a human in solving the “cocktail party” problem.

### Keywords

attention; neuroimaging; fMRI; MEG; EEG

## 1 Introduction

“How do we recognize what one person is saying when others are speaking at the same time?” With this question, E. Colin Cherry defined the “Cocktail Party Problem” six decades ago (Cherry, 1953). Attention often requires a process of selection (Carrasco, 2011). Selection is necessary because there are distinct limits on our capacity to process incoming sensory information, resulting in constant competition between inner goals and external demands (Corbetta et al., 2008). For example, eavesdropping on a particular conversation in a crowded restaurant requires top-down attention, but as soon as a baby starts to cry, this salient stimulus captures our attention automatically, due to bottom-up processing. The fact

© 2013 Elsevier B.V. All rights reserved.

Corresponding author: Adrian KC Lee, University of Washington, 1715 NE Columbia Road, Box 357988, Seattle, WA 98195, Phone: +1 (206) 616-0102, Fax: +1 (206) 221-6472, akclee@uw.edu.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

that sufficiently salient stimuli can break through our attentional focus demonstrates that all sound is processed to some degree, even when not the focus of volitional attention; however, the stimulus that is selected, whether through top-down or bottom-up control, is processed in greater detail, requiring central resources that are limited (Desimone et al., 1995). In order to operate effectively in such environments, one must be able to i) select objects of interest based on their features (e.g., spatial location, pitch) and ii) be flexible in maintaining attention on and switching attention between objects as behavioral priorities and/or acoustic scenes change. In vision research, there is a large body of work documenting the competitive interaction between volitional, top-down control and automatic, bottom-up enhancement of salient stimuli (Knudsen, 2007). However, there are comparatively fewer studies investigating how object-based auditory attention operates in complex acoustic scenes (Shinn-Cunningham, 2008). By utilizing different human neuroimaging techniques, we are beginning to understand the cortical dynamics associated with directing and redirecting auditory attention.

This review begins by providing a brief overview of neuroimaging approaches commonly used in auditory attention studies. Particular emphasis is placed on functional magnetic resonance imaging (fMRI), magnetoencephalography (MEG) and electroencephalography (EEG) because these modalities are currently used more often than other non-invasive imaging techniques, such as positron emission tomography (PET) or near-infrared spectroscopy (NIRS). To facilitate a fuller understanding of the array of neuroimaging studies, we discuss the strengths and limitations of each imaging technique as well as the ways in which the technique employed can influence how the results can be interpreted. We then review evidence that attention modulates cortical responses both in and beyond early auditory cortical areas (for a review of auditory cortex anatomy, see Da Costa et al., 2011; Woods et al., 2010). There are many models to describe auditory attention, including phenomenological models (e.g., Näätänen, 1990), which accounts for attention and automaticity in sensory organization while focusing on human neuroelectric data), behavioral models (e.g., Cowan, 1988), and neurobiological models (e.g., McLachlan et al., 2010). Many processes, from organizing the auditory scene into perceptual objects to dividing attention across multiple talkers in a crowded environment, influence auditory attention. These processes are discussed in a recent comprehensive review (Fritz et al., 2007). Here, we focus on selective attention, which Cherry cites as the key issue in allowing us to communicate in crowded cocktail parties. Moreover, we use as an organizing hypothesis that all forms of selective attention operate on perceptual objects, so in this review we focus on object-based attention (see Shinn-Cunningham, 2008 for review). This also enables us to compare and contrast results with those from the visual attention literature. We conclude by highlighting other important questions in the field of auditory attention and neuroimaging.

## 2 Methodological approaches

### 2.1 Spatial and temporal resolution considerations

Magneto- and electroencephalography (MEG, EEG; M/EEG when combined) record extracranial magnetic fields and scalp potentials that are thought to reflect synchronous post-synaptic current flow in large numbers of neurons (Hämäläinen et al., 1993). Both technologies can detect activity on the millisecond time scale characteristic of communication between neurons; the typical sampling frequency (~1,000 Hz) makes it particularly suited to studies of auditory processing, given the importance of temporal information in the auditory modality. There are important differences between MEG and EEG. For example, the skull and scalp distort magnetic fields less than electric fields, so that MEG signals are often more robust than the corresponding EEG signals. MEG is also mainly sensitive to neural sources oriented tangentially to the skull, whereas EEG is

sensitive to both radially and tangentially oriented neural sources. When MEG and EEG are used simultaneously, they can provide additional complementary information about the underlying cortical activities (Ahlfors et al., 2010; Goldenholz et al., 2009; Sharon et al., 2007). By using anatomical information obtained from magnetic resonance imaging (MRI) to constrain estimates of neural sources of observed activity, reasonable spatial resolution of cortical source can be achieved (Lee et al., 2012).

Functional MRI is another widely used non-invasive neuroimaging technique. It measures the blood-oxygenation level dependent (BOLD) signal, which reflects local changes in oxygen consumption. This BOLD signal is used as a proxy for neural activity in a particular cortical (or subcortical) location; this assumption is supported by the fact that the BOLD signal correlates strongly with the underlying local field potential in many cases (Ekstrom, 2010; Logothetis, 2008). Compared to M/EEG, fMRI has much better spatial resolution (better by a factor of about 2 to 3) but poorer temporal resolution (worse by about a factor of 1000, due to the temporal sluggishness of the BOLD signal). Figure 1 provides a summary of the tradeoffs between spatial and temporal resolution for these neuroimaging approaches.

In designing experiments, the scientific question being asked should inform the choice of which neuroimaging technique to use. For example, due to its superior spatial resolution, fMRI is well suited for a study to tease apart precisely what anatomical regions are engaged in particular tasks (e.g., comparisons of “what”/“where” processing within auditory cortical areas); in contrast, M/EEG can tease apart the dynamics of cortical activity (e.g., to temporally distinguish neural activity associated with top-down control signals before a sound stimulus begins from the signals effecting selective attention when the stimulus is playing). Other factors, apart from considerations of spatial and temporal resolution, also influence both the choice of neuroimaging technique to use and the way to interpret obtained results. These factors are summarized below.

## 2.2 Other tradeoffs related to attention studies in different techniques

**2.2.1 fMRI scanner noise**—In order to achieve good spatial and temporal resolution along with high signal-to-noise ratios, MRI scanners need powerful magnetic fields and fast switching of magnetic gradients. When a current is passed through coils inside the MRI scanner to set up these gradients, the resulting Lorentz forces cause them to vibrate, generating acoustic noise that can exceed 110 dB SPL (Counter et al., 2000; Hamaguchi et al., 2011). This scanner noise is part of the auditory scene that a subject hears during an fMRI study (Mathiak et al., 2002). As a result, in auditory paradigms involving attentional manipulation, brain activity will reflect not only activity in response to the controlled auditory stimuli, but also in response to the scanner noise, e.g., inducing involuntary orienting (Novitski et al., 2001). Sparse temporal sampling (Hall et al., 1999), wherein the stimulus is presented during silent periods between imaging acquisition, is commonly used to reduce the influence of scanner noise on the brain activity being measured. However, this technique significantly reduces the number of imaging volumes that can be acquired in a given experiment, which lowers the signal-to-noise ratio compared to continuous scanning (Huang et al., 2012). A sparse sampling strategy also decreases the temporal resolution of the measured signal acquired, making it much more difficult to estimate BOLD time courses. The sparse sampling technique does not eliminate scanner noise; it only controls the timing of the noise. Thus, the scanner noise still interacts with the controlled sound stimuli. For example, in an fMRI streaming experiment using sparse sampling, scanner noise contributed to an abnormal streaming build-up pattern (Cusack, 2005). This is consistent with the observation that auditory attention influences the formation of auditory streams (Cusack et al., 2004). Furthermore, the infrequency of the scanner noise in a sparse sampling protocol can potentially be more distracting than continuous scanning noise because the

sparse-sampling sound onsets likely trigger strong bottom-up processing in the attentional network. These factors illustrate why experimenters need to consider the psychoacoustical impact of scanner noise when employing fMRI to study auditory attention.

**2.2.2 Interpreting spatial distribution of EEG and MEG data**—Due to the ill-posed bioelectromagnetic inverse problem (Helmholtz, 1853), there is no unique solution for calculating and localizing the activity of underlying neural sources from the measured electric potential or magnetic field data. This is particularly problematic when trying to localize auditory responses. Historically, there have been debates as to whether the neural responses to auditory stimuli observed in EEG originate in frontal association areas—it took careful analysis of both EEG and MEG data to conclude that the neural generators of this response are in and around the primary auditory cortices (Hari et al., 1980). Measurements of the distribution of event-related potentials and fields (ERP/F) are useful for testing specific hypotheses related to auditory attention and many other cognitive tasks, and statistical techniques (e.g., topographical ANOVA; Murray et al., 2008) have been developed to analyze ERP/F at the sensor data level (without using neural source analysis methods, see below). However, interpretation of field topographical patterns to infer the anatomical and physiological origins requires care and judgment due to the fundamental ill-posed nature of the bioelectromagnetic inverse problem.

**2.2.3 EEG and MEG neural source analysis**—Fortunately, the inverse problem can be reformulated as a modeling problem; typically, a best solution exists, given proper regularization (Baillet, 2010). A crucial step in this modeling problem is to account for how the neural sources in the cortex are related to the M/EEG surface measurements in the presence of other tissues (e.g., scalp, skull and the brain), including the noise characteristics of this mapping. Some auditory attention studies assume generic spherical head models; others use anatomical constraints from MRI scans and boundary element methods to build individualized head models that take into account the unique geometry of the scalp, skull, and brain structure of a particular listener. The accuracy of the neural source estimation depends critically on the sensors (i.e., MEG or EEG or combined M/EEG) and the head models used.

Generally, there are two approaches in estimating the neural sources based on surface measurement: 1) source localization and 2) inverse imaging. As should be expected, the choice of what inverse estimation method to employ involves tradeoffs.

In the source localization approach, a limited number of equivalent current dipoles (ECD) can be computed based on an ordinary least-square criterion; the individual investigator must decide how many dipoles to fit. This seemingly subtle methodological decision directly, and potentially profoundly, affects the interpretation of the results. Dipole localization enables investigators to model equivalent dipoles in distinct locations (in some cases, of their choosing) on the cortex. The estimates of activity at the modeled dipoles is then fit to the observed measurements; the goodness of the model fit is typically assessed by determining what percentage of the variation in the observations can be accounted for by the estimated activity in the modeled dipoles. This approach is often favored when the investigator wants to make inferences about how activity at one site differs across experimental conditions, or to determine which subdivisions of the auditory cortices are engaged in a particular task. One advantage of ECD modeling over inverse imaging approaches is that it requires less computational power. However, it suffers from one major disadvantage: if there are any sources of neural activity other than the modeled dipoles, they can change the estimated activity at modeled locations, leading to erroneous interpretations of results. This is a particularly important point when studying auditory attention, as there is an abundance of evidence showing that solving the cocktail party problem engages a

distributed cortical network. Still, many experimentalists appear to believe that it is parsimonious and therefore best practice to account for observed measurements by solving for a small number of equivalent current dipoles in and around bilateral auditory cortex.

An alternative approach to localizing activity to brain sources is the inverse imaging method. Mathematically, this technique is based on a regularized least-squares approach. Probabilistic “priors” (a priori assumptions) are used to define the goodness of each possible solution to the under-constrained inverse problem, mapping sensor data to neural sources; the output solution is optimal, given the mathematical priors selected. Of course, the choice of these priors directly influences the estimates of source activity in a manner that is loosely analogous to the influence of the choice of how many dipoles to fit (and where they are located) on dipole localization results. Critically, however, the mathematical priors used in inverse imaging typically do not make explicit assumptions about what brain regions are involved in a given task, but instead invoke more general constraints, such as accounting for the noise characteristics of the measurements or incorporating functional priors from other observations, such as fMRI. On the negative side, however, inverse imaging approaches often require more computational power than dipole modeling, and can require acquisition of additional structural MR and coregistration data. The minimum-norm model (Hämäläinen et al., 1993), one popular inverse imaging choice, produces resultant current estimates that must necessarily be distributed in space. Other inverse methods, such as minimum-current estimates, favor sparse source estimates over solutions with many low-level, correlated sources of activity (Gramfort et al., 2012; Uutela et al., 1999).

In sum, each imaging technique has its own strengths and weaknesses. An awareness of the different assumptions that are made to estimate neural source activity from observed sensor data can help to guide the selection of whatever technique is most appropriate for a given experimental question, as well as resolve apparent discrepancies across different studies, especially when comparing interpretations of what cortical regions participate in auditory selective attention tasks based on M/EEG.

### 3 Experimental designs for neuroimaging analysis

The Posner cueing paradigm is a seminal procedure used in the study of visual attention (Posner, 1980). A central cue is used to direct endogenous attention to the most likely location of the subsequent target, while a brief peripheral cue adjacent to the subsequent target location can cue exogenous attention. By manipulating the probability that the target comes from the endogenously cued location, this paradigm allows us to compare performance when attention is directed to the target location (e.g., validly cued trials in a block where the cue is 75% likely to be valid), away from that location (e.g., invalidly cued trials in a block where the cue is 75% likely to be valid), or distributed across all locations (trials in a block where the cue is 50% likely to be valid; e.g., see Carrasco, 2011).

Variants of this paradigm have been used to study auditory attention, typically using simple stimuli with one auditory object presented at a time much like the visual stimuli used in past versions of the task (Roberts et al., 2006; Roberts et al., 2009). For example, using simple stimuli can be an efficient way to investigate attention based on different feature cues (such as location or pitch). However, it has been suggested that in order to understand the neural substrate of attention, the auditory system should be placed under high load conditions (Hill et al., 2010). Furthermore, assuming that the spectrotemporal elements are perceptually segregated into distinct objects, a competing object must be presented simultaneously with the target to study what brain areas allow a listener to select the correct object in a scene.

In most auditory attention tasks using a Posner-like procedure, the listener is cued to attend to the feature of an upcoming stimulus. During the stimulus interval, he or she must make a



judgment about the object in the scene that has the cued feature; finally, the listener responds to ensure that the listener performed the task correctly (Figure 2). In order to separate motor activities related to the response from cortical activity related to the attentional task, the response period is generally delayed in both fMRI and M/EEG paradigms. Some fMRI procedures include catch trials, in which the cue is presented but no stimuli are presented afterwards, to isolate activity in the pre-stimulus preparatory period and determine which cortical regions are engaged in preparing to attend to a source with a particular attribute, rather than evoked by the attended stimulus (e.g., Hill et al., 2010). However, unlike M/EEG, the slow dynamics of the BOLD response make it impossible to use fMRI to determine the time course of cortical activity related to directing attention, or to completely isolate such activity from activity involved in selecting the desired auditory object. Even in M/EEG studies, where pre-auditory-stimulus and post-auditory-stimulus responses are temporally distinct, the choice of the baseline-correction period can influence interpretation of these measures. For instance, consider the hypothetical case illustrated in Figure 2. If baseline is defined by the pre-stimulus period, as opposed to pre-cue period (hatched vs. solid gray bars), one could erroneously conclude that attention does not play a role in this example (see Urbach et al., 2006 for a detailed discussion).

Finally, it goes without saying that behavioral paradigm should always be carefully designed in order to highlight attention-driven cortical activity using the applied neuroimaging methods. For example, using identical stimuli across different attentional conditions ensures that any neural differences observed come from differences in the attentional goals of the listener, rather than due to stimulus differences. These sorts of issues should be considered when interpreting results of past auditory attention studies, and when designing paradigms for future investigations.

## 4 Drawing from theories of visual attention

The body of literature describing neural mechanisms of selective visual attention is much richer than the literature on selective auditory attention. The attentional network is often assumed to be supramodal—not specific to one sensory modality (Knudsen, 2007)—but the nodes in this network have been primarily mapped out using visual experiments (Figure 3; Corbetta et al., 2008). It is thus important to ask questions about auditory attention with a deep appreciation of the visual attention literature, all the while being cognizant of the intrinsic differences between these two sensory modalities. Just as in vision, object formation and object selection are critical to selective auditory attention, providing us with a common framework to discuss attention across sensory modality (Shinn-Cunningham, 2008). In this section, we briefly discuss how theories of object-based visual attention have influenced our conceptualization of auditory attention.

### 4.1 Spatial and non-spatial feature attention in the “what”/“where” pathways

We can direct our attention to an auditory object based on either spatial or non-spatial features. The dual-pathway theory for feature and spatial processing has long been accepted in vision (Desimone et al., 1995; Mishkin et al., 1983). A similar dual pathway may control auditory information processing, with a posterior parietal pathway (i.e., a postero-dorsal stream) subserving spatial processing and a temporal pathway (i.e., an antero-ventral stream) handling identification of complex objects (see Rauschecker et al., 2009 for a review). Furthermore, evidence from both non-human as well as human primates suggests that the primary auditory cortex and surrounding areas are also organized spatially to feed into this dual-pathway configuration (Rauschecker, 1998). This conceptualization has inspired many auditory neuroimaging studies to focus on contrasting the differences in cortical processing based on spatial and non-spatial features using a wide variety of behavioral tasks. In the

present review, we thus highlight studies that focus on differences in cortical activity when attention is based on spatial vs. non-spatial acoustic features.

#### 4.2 Attentional network – role of FEF, IPS and TPJ

In vision, spatial attention and eye gaze circuitry are intimately linked because sensory acuity changes with eccentricity from the fovea. Two important cortical nodes that participate both in visual attention and gaze control are the frontal eye fields region (FEF) and the intraparietal sulcus (IPS). The FEF, located in the premotor cortex, controls eye gaze but also participates in directing spatial attention independent of eye movement, i.e. covert attention (Bruce et al., 1985; Wardak et al., 2006). In audition, FEF activity has been shown to represent sounds both in retinal and extra-retinal space (Tark et al., 2009). Other visual fMRI studies show that the IPS contains multiple retinotopic spatial maps that are engaged by attention (Szczepanski et al., 2010; Yantis, 2008). Moreover, disrupting IPS using transcranial magnetic stimulation (TMS) can alter visual spatial attention (Szczepanski et al., 2013). Auditory neuroimaging studies implicate IPS as playing an automatic, stimulus-driven role in figure-ground segregation (Cusack, 2005; Teki et al., 2011) as well as participating in auditory spatial working memory tasks (Alain et al., 2010). Finally, the temporoparietal junction (TPJ) is thought to contribute to the human auditory P3 response (Knight et al., 1989). Right TPJ (rTPJ), specifically, is thought to act as a “filter” of incoming stimuli, sending a “circuit-breaking” or interrupt signal that endogenously changes the locus of attention to an important stimulus outside of the current focus of attention (Corbetta et al., 2008). rTPJ registers salient events in the environment not only in the visual, but also in the auditory and tactile modalities (Downar et al., 2000).

The roles that FEF, IPS, and rTPJ play in controlling attention have been studied primarily in vision, which is a spatiocentric modality. Past fMRI studies suggest that a frontoparietal network (including FEF and IPS) is engaged in directing top-down attention to both spatial and non-spatial features (Giesbrecht et al., 2003). But even during this pre-stimulus interval, areas of the visual cortex selective for the to-be-attended features are also activated (Giesbrecht et al., 2006; Slagter et al., 2006). This pre-stimulus activity is similar to the contralateral activation in Heschl’s gyrus (HG) and planum temporale (PT) observed when a listener is preparing to attend to a sound from a given direction (Voisin, 2006). In Section 6.1, we review evidence for how this network functions when both spatial and non-spatial features are used to focus auditory attention.

### 5 Evidence of attentional modulation in auditory cortex

Many studies have examined the role of the auditory cortical sensory areas in solving the cocktail party problem. Early studies specifically addressed whether attention has a modulatory effect on early cortical responses in different selective attention tasks. In one groundbreaking ERP study (Hillyard et al., 1973), listeners responded to a deviant tone while selectively attending to one of the two streams of tone pips, with one ear presented with the lower frequency and the other, the higher frequency. Using only one electrode at the vertex of the head, researchers found that the attentional state of the listener strongly modulated the N100. Attention also modulates the magnetic counterpart of this component, the M100 (Woldorff et al., 1993). Using ECD modeling in conjunction with MRI, this M100 component was localized to the auditory cortex on the supratemporal plane, just lateral to HG. There was also a significant attention modulatory effect on earlier components of the magnetic response (M20-50), suggesting that top-down mechanisms bias the response to an auditory stimulus at (and possibly before) the initial stages of cortical analysis to accomplish attentional selection. Subsequent fMRI studies using different behavioral paradigms also showed attentional enhancement of BOLD activity in the primary and secondary auditory cortices (e.g., Grady et al., 1997; Hall et al., 2000; Petkov et al., 2004; Rinne et al., 2005).

Specifically, recent evidence suggests that there is a frequency-specific, attention-specific enhancement in the response of the primary auditory cortex, without any response suppression for the unattended frequency. There also seems to be more widespread general enhancement across auditory cortex when performing an attentional task compared to passive listening (Paltoglou et al., 2009; Paltoglou et al., 2011). Furthermore, in dichotic listening tasks, HG and planum polare (PP) were more active on the hemisphere contralateral to side of the attended auditory source than on the ipsilateral side (e.g., Ciaramitaro et al., 2007; Jancke et al., 2002; Rinne, 2010; Rinne et al., 2008; Yang et al., 2013).

Evidence also suggests that attention based on object-related (“what”) and spatial (“where”) features modulates different subdivisions of the auditory cortex, similar to the dual stream functional organization in vision (Rauschecker, 1998). Sub-regions of PT and HG are more active depending on whether listeners are judging a sequence of sounds that change either in pitch or location token by token, lending support to the view that there are distinct cortical areas for processing spatial and object properties of complex sounds (Warren et al., 2003); however, it should be noted that the stimuli used in this study were not identical across conditions, which complicates interpretation). More recently, in a multimodal imaging study (using fMRI-weighted MEG inverse imaging as well as MEG ECD analysis), listeners attended to either the phonetic or the spatial attribute of a pair of tokens and responded when a particular value of that feature was repeated (Ahveninen et al., 2006). Leveraging the strengths of both imaging modalities, the authors found an anterior “what” and posterior “where” pathway, with “where” activation leading “what” by approximately 30 ms. This finding suggests that attention based on different features can modulate distinct local neuronal networks dynamically based on situational requirements.

The aforementioned studies used brief stimuli to assess auditory attention, but in everyday listening situations, we must often selectively attend to an ongoing stream of sound. Understanding the dynamics of auditory attention thus requires experimental designs with streams of sound and different analyses. Luckily, MEG and EEG are particularly suited to measure continuous responses. Indeed, evidence from studies in which listeners heard one stream suggests that MEG and EEG signals track the slow (2–20 Hz) acoustic envelope of an ongoing stimulus (Abrams et al., 2008; Ahissar et al., 2001; Aiken et al., 2008). These single-stream findings can be extended to studies of selective attention involving more than one stream. In one recent study, listeners attended to one of two polyrhythmic tone sequences and tried to detect a rhythmic deviant in either the slow (4 Hz) or the fast (7 Hz) isochronous stream (Xiang et al., 2010). The neural power spectral density (obtained from a subset of the strongest MEG channels) showed an enhancement at the neural signal at the repetition frequency of the attended target. This frequency-specific attentional modulation was accompanied by an increase in long-distance coherence across sensors. The source of the increase in spectral power was localized to the auditory cortex (HG and superior temporal gyrus), suggesting that attention shapes responses in this cortical region.

One disadvantage of studying attention by using competing rhythmic streams with different repetition rates is that it is intrinsically tied to the frequencies used in the stimuli. Furthermore, temporal resolution is lost when the steady-state spectral content of the signal is analyzed, collapsed across time. In order to allow acquisition of temporally detailed responses to continuous auditory stimuli (e.g., speech), a technique known as the AESPA (auditory evoked spread spectrum analysis) was developed wherein the auditory cortical response to the time-varying continuous stimuli is estimated based on the measured neural data (Lalor et al., 2010; Lalor et al., 2009). This technique was used to analyze EEG signals recorded when listeners selectively attended to one of two spoken stories (Power et al., 2012). Questions probing comprehension of the target story were used to verify the



listener's attentional engagement. The AESPA responses derived from the neural data showed a robust, left-lateralized attention effect peaking at ~200 ms. These results confirm that attention modulates cortical responses. However, this method is insensitive to any cortical activity unrelated to the stimulus envelope and assumes a simple linear relationship between the stimulus envelope and the EEG response. Future studies need to further investigate how AESPA responses are related to cortical attentional modulation.

A number of other recent MEG and EEG studies have also addressed the question of how selectively attending to one ongoing stream in a multi-stream environment affects neural signals. In a study where listeners attended to one of two melodic streams at distinct spatial locations, EEG responses to note onsets in the attended stream were 10 dB higher than responses to unattended note onsets (Choi et al., 2013). These differences in neural responses to attended versus unattended streams were large enough that a single EEG response to a 3 s long trial could be used to reliably classify which stream the listener was attending. In another study, listeners selectively attended to one speech stream in a dichotic listening task, and EEG recordings showed a similar gain-control-like effect in or near HG, based on a template matching procedure using N100-derived source waveforms (Kerlin et al., 2010). Two other MEG studies—one in which listeners selectively attended to a target speech stream in the presence of a background stream separated in space (Ding et al., 2012a), and the other in which listeners selectively attended to a male speaker or a female speaker, presented diotically (Ding et al., 2012b) support a model of attention-modulated gain control, wherein attention serves to selectively enhance the response to the attended stream. These studies further found that the neural representation of the target speech stream is insensitive to an intensity change in either the target or the background stream, suggesting that this attentional gain does not apply globally to all streams in the auditory scene, but rather modulates responses in an object-specific manner.

So far we have discussed studies where listeners selected a stream in the presence of other auditory streams. When listeners are presented with audio-visual stimuli and told to selectively attend to one modality at a time, activity in the lateral regions of the auditory cortex is modulated (Petkov et al., 2004; Woods et al., 2009). This is in contrast to the more medial primary regions of auditory cortex, which are typically modulated more by the characteristics of the auditory stimulus. Furthermore, visual input plays an important role especially in speech processing in a multi-talker environment. When listeners are selectively watching a speaker's face, the auditory cortex can better track the temporal speech envelope of the attended speaker compared to when the speaker's face is not available (Zion Golumbic et al., 2013), and regions of interest in auditory cortex differentially sensitive to high- and low-frequency tones modulate the strength of their responses based on whether the frequency of the currently attended set of tones (Da Costa et al., 2013).

Taken together, results from these studies suggest that the neural representation of sensory information is modulated by attention in the human auditory cortex, consistent with recent findings obtained in intracranial recordings (Mesgarani et al., 2012). However, many of these studies employed analyses that either implicitly or explicitly assumed that attention modulates only early stage processing in the auditory sensory cortex, e.g., by fitting source waveforms based on the early components of an ERP/F response (N100/M100) or by using a region-of-interest approach that analyzes only primary auditory cortical areas.

## 6 Attentional modulation beyond the auditory cortex

Studies described in the previous section primarily focus on how attention modulates stimulus coding in the auditory cortex while listeners perform an auditory task. However, in a multi-talker environment (e.g., in a conference poster session), you are not always actively

attending to one sound stream. For instance, you may be preparing to pick out a conversation (e.g., covertly monitoring your student who is about to present a poster) or switching attention to a particularly salient stimulus (e.g., your program officer just called your name). In this section, we will present studies that investigate the cortical responses associated with directing, switching, and maintaining auditory attention.

### 6.1 Directing spatial and non-spatial feature-based auditory attention

As discussed in Sections 4.1 and 5, there is evidence that dual pathways exist in auditory cortex that preferentially encode spatial and non-spatial features of acoustic inputs, paralleling the “what” and “where” pathways that encode visual inputs (Desimone et al., 1995; Mishkin et al., 1983; Rauschecker et al., 2009). This raises the question of how endogenous direction of auditory attention results in top-down preparatory signals even before an acoustic stimulus begins, comparable to the kinds of preparatory signals observed when visual attention is deployed.

A number of studies have explored whether the preparatory signals engaging attention arise from the same or different control areas, depending on whether attention is directed to a spatial or non-spatial feature. Left and right premotor and inferior parietal areas show an increased BOLD response when listeners are engaged in a sound localization task compared to a sound recognition task (Mayer et al., 2006) and bilateral Brodmann Area 6 (most likely containing FEF; (Rosano et al., 2002) is more active when attending stimuli based on location rather than pitch (Degerman et al., 2006), suggesting different control areas for spatial versus feature-based attention. It is often hard to differentiate between the neural activity associated with a listener preparing to attend to a particular auditory stimulus versus the activity evoked during auditory object selection, especially in fMRI studies (with their limited temporal resolution; see also Section 2.1). However, in a recent fMRI study using catch trials to isolate the preparatory control signals associated with directing spatial and pitch-based attention, both bilateral premotor and parietal regions were more active for an upcoming location trial than a pitch trial; in contrast, left inferior frontal gyrus (linked to language processing) was more active during preparation to attend to source pitch compared to source location (Hill et al., 2010). In other words, these regions are implicated in controlling attention given that they were active in trials where listeners were preparing to attend, even when no stimulus is ultimately presented. Further, it appears that attention to different features (e.g., space or pitch) involves control from distinct areas. Also in this study, bilateral superior temporal sulci (STS) were also found to be more active during the stimulus period when listeners were attending to the pitch of the stimulus rather than the location. However, the sluggishness of the BOLD response obscures any underlying rapid cortical dynamics. A more recent MEG study (inverse imaging using a distributed model constrained with MRI information) found that left FEF activity is enhanced in preparation and during a spatial attention task while the left posterior STS (previously implicated for pitch categorization) is greater in preparation for a pitch attention task (Lee et al., 2013). These findings are in line with the growing number of neuroimaging studies (e.g., Smith et al., 2010, fMRI; Diaconescu et al., 2011, EEG) suggesting that control of auditory spatial attention engages cortical circuitry that is similar to that engaged during visual attention—though the exact neuronal subpopulations controlling visual and auditory attention within a region may be distinct (Kong et al., 2012); in contrast, selective attention to a non-spatial feature seems to invoke activity in sub-networks specific to the attended sensory modality.

### 6.2 Switching attention between auditory streams

Similar to the challenges of isolating neural activities associated with directing top-down attention versus auditory object selection, it is also difficult to tease apart top-down and bottom-up related processes when studying the switching of attention. Nonetheless,

converging evidence suggests that the frontoparietal network associated with visual attention orientation also participates in auditory switching of attention. When listeners switched attention between a male and female stream (both streams presented diotically, with attention switching cued by a spoken keyword), activation in the posterior parietal cortex (PPC) is higher compared to when listeners maintain fixation throughout the trial (Shomstein et al., 2006). Similarly, when listeners switch attention from one ear to the other (with male and female streams presented monaurally to each ear), a similar region in the PPC is activated, suggesting that PPC participates in the control of auditory attention, similar to its role in similar visual attention switching. It thus appears that PPC participates in both spatial and nonspatial attention in the auditory and visual modalities (Serences et al., 2004; Yantis et al., 2002). The parietal cortices also have been implicated in involuntary attention switching, exhibiting “change detection” responses in paradigms using repeated “standard” sounds with infrequent “oddballs” (Molholm et al., 2005; Watkins et al., 2007).

One study employing fMRI used a similar auditory and visual attention switching task to compare and contrast the orienting and maintenance of spatial attention in audition and vision (Salmi et al., 2007b; but also see related ERP study, Salmi et al., 2007a). In a conjunction analysis, they showed that parietal regions are involved in orienting spatial attention for both modalities. Interestingly, they found that orienting-related activity in the TPJ was stronger in the right hemisphere than in the left in the auditory task, but no such asymmetry was found in the visual task. Similarly, using a Posner-like cueing paradigm, another fMRI study examined the switching of auditory attention (Mayer et al., 2009). Listeners were cued by a monaural auditory tone in the same ear as the target stimulus. In one block, these cues were informative (75% valid) in order to promote top-down orientation of spatial attention; in a contrasting block, these cues were uninformative (50% valid). Precentral areas (encompassing FEF) and the insula were more active during uninformative cue trials, suggesting that these areas are associated with automatic orienting of attention. However, a later paper argued that the Mayer study did not separate the processes of auditory cued attention shifting from target identification (Huang et al., 2012). This later study found that voluntary attention switching and target discrimination both activate a frontoinsula-cingular attentional network that includes the anterior insula, inferior frontal cortex, and medial frontal cortices (consistent with another fMRI study on cue-guided attention shifts; (Salmi et al., 2009). In agreement with other studies, this study also reported that cued spatial-attention shifting engaged bilateral precentral/FEF regions (e.g., Garg et al., 2007) and posterior parietal areas (e.g., Shomstein et al., 2006); the study also showed that the right FEF was engaged by distracting events that catch bottom-up attention.

Some of the results presented above may have been confounded by the presence of acoustical scanner noise. There are not many fMRI studies of auditory switching that use a sparse sampling design (except see Huang et al., 2012). By comparison, it is much easier to control for the acoustical environment in M/EEG studies. This, coupled with the excellent temporal resolution of M/EEG, has helped build the extensive literature documenting the mismatch negativity (MMN), novelty-P3, and reorienting negativity (RON) response markers for different stages of involuntary auditory attention shifting (see Escera et al., 2007). As discussed in Section 2.2.2, ERP/F analyses are often difficult to interpret in conjunction with fMRI studies because it is non-trivial to relate ERP/F spatial topography to specific cortical structures. A recent study using dipole modeling to ascertain the neural generators of response components related to switching of attention found that, consistent with many previous studies (e.g., Rinne et al., 2000), N1/MMN components were localized in the neighborhood of supratemporal cortex, while the RON is localized to the left precentral sulcus/FEF area (Horváth et al., 2008). However, as in previous studies (e.g., Alho et al., 1998), they were unable to localize the P3 component using the ECD method,

likely because there was more than a single equivalent current dipole contributing to this temporal component of the neural response.

A recent M/EEG study (using inverse imaging combined with MRI information) looked at the brain dynamics associated with switching spatial attention in response to a visual cue (Larson et al., 2013). They found that rTPJ, rFEF, and rMFG were more active when listeners were prompted by a visual cue to switch spatial attention immediately prior to the auditory target interval compared to when they maintained attention to the originally cued hemifield. Furthermore, the normalized behavioral performance difference in switch- versus maintain-attention conditions was correlated with the normalized rTPJ and rFEF activation differences across subjects (Figure 4). The recruitment of rFEF in reaction to a switch visual cue (1/3 of the trials) compared to a maintain attention cue (2/3 of the trials) are both consistent with the finding that the rFEF is activated by distracting events that catch attention in a bottom-up manner (Huang et al., 2012; Salmi et al., 2009). This study provides further evidence that rTPJ, rFEF, and rMFG operate within a supramodal cortical attention network, and that these areas are directly involved with successful switching of auditory attention.

Taken as a whole, switching attention in the auditory domain recruits a similar cortical network engaged by switching attention between visual objects (Corbetta et al., 2008), providing further evidence that the attentional network is supramodal (Knudsen, 2007). However, most of these auditory paradigms involve switching spatial attention only. It remains to be seen whether switching of non-spatial feature-based auditory attention would recruit other specialized auditory or task-relevant regions.

## 7 Conclusions and future directions

Functional neuroimaging techniques such as fMRI, MEG, and EEG, can be used to investigate the brain's activity during tasks that require the deployment of attention. Using carefully designed paradigms that heed each imaging technology's inherent tradeoffs, we have learned much about the modulatory effects of attention on sensory processing. In particular, we find that attention to an auditory stimulus alters the representation of sounds in the auditory cortex, especially in secondary areas. Auditory attention modulates activity in other parts of the cortex as well, including canonical areas implicated in control of visual attention, where they have been studied in detail. The fact that attention in different modalities engages similar regions makes sense given the hypothesized conservation of the attentional networks controlling visual and auditory attention. Studies show that the attentional control network is most likely supramodal, but is deployed differently depending on the sensory modality attended, as well as whether attention is based on space or some non-spatial feature like pitch. More work in this area will lead to a better understanding of how these networks operate, including the extent to which different areas participate unimodally or depending on particular task parameters.

While it is useful to map out regions of the cortex participating in different aspects of auditory attention, an important question still remains: how are these areas functionally coupled to the auditory cortex (Figure 3)? An important future direction is to take the systems neuroscience perspective and address attention's potential modulation of the brain's connectivity (Banerjee et al., 2011). The use of TMS, in conjunction with other neuroimaging modalities (e.g., fMRI, EEG), can also provide important information about the causal interactions between these cortical nodes participating in auditory attention. PET and NIRS may also play an important role in mapping auditory attention in cochlear implantees (Ruytjens et al., 2006) and infants (Wilcox et al., 2012), respectively. Work on mitigating the noise associated with fMRI will also allow for more diverse experimental

designs (Peelle et al., 2010). Finally, while recent evidence from electrocorticographical studies reveals the importance of ongoing oscillatory activity in auditory selective attention more directly than is possible with more standard imaging methods (e.g., Lakatos et al., 2013), non-invasive M/EEG remains an important and convenient tool for studying the role that cortical rhythms play in segregating and selecting a source to be attended at a cocktail party.

## Acknowledgments

Thanks to Frederick J. Gallun for suggestions. This work was supported by National Institute on Deafness and Other Communication Disorders grant R00DC010196 (AKCL), training grants T32DC000018 (EL), T32DC005361 (RKM), fellowship F32DC012456 (EL), CELEST, an NSF Science of Learning Center (SBE-0354378; BGSC), and an NSSEFF fellowship to BGSC.

## Abbreviations

<b>AESPA</b>	auditory evoked spread spectrum analysis
<b>BOLD</b>	blood-oxygenation level dependent
<b>ECD</b>	equivalent current dipole
<b>EEG</b>	electroencephalography
<b>ERF</b>	event-related field
<b>ERP</b>	event-related potential
<b>FEF</b>	frontal eye fields
<b>(f)MRI</b>	(functional) magnetic resonance imaging
<b>HG</b>	Heschl's gyrus
<b>IPS</b>	intraparietal sulcus
<b>NIRS</b>	near-infrared spectroscopy
<b>MEG</b>	magnetoencephalography
<b>MMN</b>	mismatch negativity
<b>RON</b>	reorienting negativity
<b>PET</b>	positron emission tomography
<b>PP</b>	planum polare
<b>PPC</b>	posterior parietal cortex
<b>PT</b>	planum temporale
<b>(r)MFG</b>	(right) middle frontal gyrus
<b>(r)TPJ</b>	(right) temporoparietal junction
<b>STS</b>	superior temporal sulcus
<b>TMS</b>	transcranial magnetic stimulation

## References

- Abrams DA, Nicol T, Zecker S, Kraus N. Right-hemisphere auditory cortex is dominant for coding syllable patterns in speech. *Journal of Neuroscience*. 2008; 28:3958–3965. [PubMed: 18400895]
- Ahissar E, Nagarajan S, Ahissar M, Protopapas A, Mahncke H, Merzenich MM. Speech comprehension is correlated with temporal response patterns recorded from auditory cortex.



- Proceedings of the National Academy of Sciences of the United States of America. 2001; 98:13367–13372. [PubMed: 11698688]
- Ahlfors SP, Han J, Belliveau JW, Hämäläinen MS. Sensitivity of MEG and EEG to Source Orientation. *Brain Topography*. 2010; 23:227–232. [PubMed: 20640882]
- Ahveninen J, Jaaskelainen I, Raij T, Bonmassar G, Devore S, Hämäläinen M, Levanen S, Lin F, Sams M, Shinn-Cunningham B, Witzel T, Belliveau J. Task-modulated “what” and “where” pathways in human auditory cortex. *Proceedings of the National Academy of Sciences of the United States of America*. 2006; 103:14608–14613. [PubMed: 16983092]
- Aiken SJ, Picton TW. Human cortical responses to the speech envelope. *Ear and Hearing*. 2008; 29:139–157. [PubMed: 18595182]
- Alain C, Shen D, Yu H, Grady C. Dissociable Memory- and Response-Related Activity in Parietal Cortex During Auditory Spatial Working Memory. *Frontiers in Psychology*. 2010; 1
- Alho K, Winkler I, Escera C, Huottilainen M, Virtanen J, Jaaskelainen IP, Pekkonen E, Ilmoniemi RJ. Processing of novel sounds and frequency changes in the human auditory cortex: magnetoencephalographic recordings. *Psychophysiology*. 1998; 35:211–224. [PubMed: 9529947]
- Baillet, S. The Dowser in the Fields: Searching for MEG Sources. In: Hansen, P.; Kringelbach, M.; Salmelin, R., editors. *MEG: An Introduction to Methods*. Oxford University Press; Oxford: 2010. p. 83-123.
- Banerjee A, Pillai AS, Horwitz B. Using large-scale neural models to interpret connectivity measures of cortico-cortical dynamics at millisecond temporal resolution. *Frontiers in Systems Neuroscience*. 2011; 5:102. [PubMed: 22291621]
- Boynton GM, Engel SA, Glover GH, Heeger DJ. Linear systems analysis of functional magnetic resonance imaging in human V1. *Journal of Neuroscience*. 1996; 16:4207–4221. [PubMed: 8753882]
- Bruce CJ, Goldberg ME. Primate frontal eye fields. I. Single neurons discharging before saccades. *Journal of Neurophysiology*. 1985; 53:603–635. [PubMed: 3981231]
- Carrasco M. Visual attention: The past 25 years. *Vision Research*. 2011; 51:1484–1525. [PubMed: 21549742]
- Cherry E. Some Experiments on the Recognition of Speech, with One and with Two Ears. *Journal of the Acoustical Society of America*. 1953; 25:975–979.
- Choi I, Rajaram S, Varghese LA, Shinn-Cunningham BG. Quantifying attentional modulation of auditory-evoked cortical responses from single-trial electroencephalography. *Frontiers in Human Neuroscience*. 2013; 7:115. [PubMed: 23576968]
- Ciaramitaro VM, Buracas GT, Boynton GM. Spatial and Cross-Modal Attention Alter Responses to Unattended Sensory Information in Early Visual and Auditory Human Cortex. *Journal of Neurophysiology*. 2007; 98:2399–2413. [PubMed: 17715196]
- Corbetta M, Patel G, Shulman GL. The reorienting system of the human brain: from environment to theory of mind. *Neuron*. 2008; 58:306–324. [PubMed: 18466742]
- Counter SA, Olofsson A, Borg E, Bjelke S, Häggström A, Grahn HF. Analysis of magnetic resonance imaging acoustic noise generated by a 4.7 T experimental system. *Acta Oto-Laryngologica*. 2000; 120:739–743. [PubMed: 11099151]
- Cowan N. Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*. 1988; 104:163–191. [PubMed: 3054993]
- Cusack R. The intraparietal sulcus and perceptual organization. *Journal of Cognitive Neuroscience*. 2005; 17:641–651. [PubMed: 15829084]
- Cusack R, Deeks J, Aikman G, Carlyon R. Effects of location, frequency region, and time course of selective attention on auditory scene analysis. *Journal of Experimental Psychology-Human Perception and Performance*. 2004; 30:643–656. [PubMed: 15301615]
- Da Costa S, van der Zwaag W, Miller LM, Clarke S, Saenz M. Tuning In to Sound: Frequency-Selective Attentional Filter in Human Primary Auditory Cortex. *Journal of Neuroscience*. 2013; 33:1858–1863. [PubMed: 23365225]

- Da Costa S, van der Zwaag W, Marques JP, Frackowiak RSJ, Clarke S, Saenz M. Human Primary Auditory Cortex Follows the Shape of Heschl's Gyrus. *Journal of Neuroscience*. 2011; 31:14067–14075. [PubMed: 21976491]
- Degerman A, Rinne T, Salmi J, Salonen O, Alho K. Selective attention to sound location or pitch studied with fMRI. *Brain Research*. 2006; 1077:123–134. [PubMed: 16515772]
- Desimone R, Duncan J. Neural Mechanisms of Selective Visual-Attention. *Annual Review of Neuroscience*. 1995; 18:193–222.
- Diaconescu AO, Alain C, McIntosh AR. Modality-dependent “What” and “Where” Preparatory Processes in Auditory and Visual Systems. *Journal of Cognitive Neuroscience*. 2011; 23:1–15. [PubMed: 20586702]
- Ding N, Simon JZ. Neural coding of continuous speech in auditory cortex during monaural and dichotic listening. *Journal of Neurophysiology*. 2012a; 107:78–89. [PubMed: 21975452]
- Ding N, Simon JZ. Emergence of neural encoding of auditory objects while listening to competing speakers. *Proceedings of the National Academy of Sciences of the United States of America*. 2012b; 109:11854–11859. [PubMed: 22753470]
- Downar J, Crawley AP, Mikulis DJ, Davis KD. A multimodal cortical network for the detection of changes in the sensory environment. *Nature Neuroscience*. 2000; 3:277–283.
- Ekstrom A. How and when the fMRI BOLD signal relates to underlying neural activity: The danger in dissociation. *Brain Research Reviews*. 2010; 62:233–244. [PubMed: 20026191]
- Escera C, Corral M. Role of mismatch negativity and novelty-P3 in involuntary auditory attention. *Journal of Psychophysiology*. 2007; 21:251.
- Fritz JB, Elhilali M, David SV, Shamma SA. Auditory attention—focusing the searchlight on sound. *Current Opinion in Neurobiology*. 2007; 17:437–455. [PubMed: 17714933]
- Garg A, Schwartz D, Stevens AA. Orienting auditory spatial attention engages frontal eye fields and medial occipital cortex in congenitally blind humans. *Neuropsychologia*. 2007; 45:2307–2321. [PubMed: 17397882]
- Giesbrecht B, Woldorff MG, Song AW, Mangun GR. Neural mechanisms of top-down control during spatial and feature attention. *NeuroImage*. 2003; 19:496–512. [PubMed: 12880783]
- Giesbrecht B, Weissman DH, Woldorff MG, Mangun GR. Pre-target activity in visual cortex predicts behavioral performance on spatial and feature attention tasks. *Brain Research*. 2006; 1080:63–72. [PubMed: 16412994]
- Goldenholz DM, Ahlfors SP, Hämäläinen MS, Sharon D, Ishitobi M, Vaina LM, Stufflebeam SM. Mapping the signal-to-noise-ratios of cortical sources in magnetoencephalography and electroencephalography. *Human Brain Mapping*. 2009; 30:1077–1086. [PubMed: 18465745]
- Grady CL, Van Meter JW, Maisog JM, Pietrini P, Krasuski J, Rauschecker JP. Attention-related modulation of activity in primary and secondary auditory cortex. *Neuroreport*. 1997; 8:2511–2516. [PubMed: 9261818]
- Gramfort A, Kowalski M, Hämäläinen M. Mixed-norm estimates for the M/EEG inverse problem using accelerated gradient methods. *Physics in Medicine and Biology*. 2012; 57:1937–1961. [PubMed: 22421459]
- Hall DA, Haggard MP, Akeroyd MA, Summerfield AQ, Palmer AR, Elliott MR, Bowtell RW. Modulation and task effects in auditory processing measured using fMRI. *Human Brain Mapping*. 2000; 10:107–119. [PubMed: 10912590]
- Hall DA, Haggard MP, Akeroyd MA, Palmer AR, Summerfield AQ, Elliott MR, Gurney EM, Bowtell RW. “Sparse” temporal sampling in auditory fMRI. *Human Brain Mapping*. 1999; 7:213–223. [PubMed: 10194620]
- Hamaguchi T, Miyati T, Ohno N, Hirano M, Hayashi N, Gabata T, Matsui O, Matsushita T, Yamamoto T, Fujiwara Y, Kimura H, Takeda H, Takehara Y. Acoustic noise transfer function in clinical MRI a multicenter analysis. *Academic radiology*. 2011; 18:101–106. [PubMed: 21145030]
- Hämäläinen M, Hari R, Ilmoniemi R, Knuutila J, Lounasmaa O. Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics*. 1993; 65:413–497.

- Hari R, Aittoniemi K, Järvinen ML, Katila T, Varpula T. Auditory evoked transient and sustained magnetic fields of the human brain. Localization of neural generators. *Experimental Brain Research*. 1980; 40:237–240. [PubMed: 7428878]
- Helmholtz, Hv. Ueber einige Gesetze der Vertheilung elektrischer Ströme in körperlichen Leitern, mit Anwendung auf die thierisch-elektrischen Versuche. *Ann Phys Chem*. 1853; 89:211–233.
- Hill KT, Miller LM. Auditory attentional control and selection during cocktail party listening. *Cerebral Cortex*. 2010; 20:583–590. [PubMed: 19574393]
- Hillyard SA, Hink RF, Schwent VL, Picton TW. Electrical signs of selective attention in the human brain. *Science*. 1973; 182:177–180. [PubMed: 4730062]
- Horváth J, Maess B, Berti S, Schröger E. Primary motor area contribution to attentional reorienting after distraction. *Neuroreport*. 2008; 19:443–446. [PubMed: 18287943]
- Huang S, Belliveau JW, Tengshe C, Ahveninen J. Brain networks of novelty-driven involuntary and cued voluntary auditory attention shifting. *PLoS ONE*. 2012; 7:e44062. [PubMed: 22937153]
- Jancke L, Shah NJ. Does dichotic listening probe temporal lobe functions? *Neurology*. 2002; 58:736–743. [PubMed: 11889237]
- Kerlin JR, Shahin AJ, Miller LM. Attentional gain control of ongoing cortical speech representations in a “cocktail party”. *Journal of Neuroscience*. 2010; 30:620–628. [PubMed: 20071526]
- Knight RT, Scabini D, Woods DL, Clayworth CC. Contributions of temporal-parietal junction to the human auditory P3. *Brain Research*. 1989; 502:109–116. [PubMed: 2819449]
- Knudsen EI. Fundamental components of attention. *Annual Review of Neuroscience*. 2007; 30:57–78.
- Kong L, Michalka SW, Rosen ML, Sheremata SL, Swisher JD, Shinn-Cunningham BG, Somers DC. Auditory spatial attention representations in the human cerebral cortex. *Cerebral Cortex*. 2012
- Lakatos P, Musacchia G, O’Connell MN, Falchier AY, Javitt DC, Schroeder CE. The Spectrotemporal Filter Mechanism of Auditory Selective Attention. *Neuron*. 2013; 77:750–761. [PubMed: 23439126]
- Lalor EC, Foxe JJ. Neural responses to uninterrupted natural speech can be extracted with precise temporal resolution. *European Journal of Neuroscience*. 2010; 31:189–193. [PubMed: 20092565]
- Lalor EC, Power AJ, Reilly RB, Foxe JJ. Resolving precise temporal processing properties of the auditory system using continuous stimuli. *Journal of Neurophysiology*. 2009; 102:349–359. [PubMed: 19439675]
- Larson E, Lee AKC. The cortical dynamics underlying effective switching of auditory spatial attention. *NeuroImage*. 2013; 64:365–370. [PubMed: 22974974]
- Lee AKC, Larson E, Maddox RK. Mapping cortical dynamics using simultaneous MEG/EEG and anatomically-constrained minimum-norm estimates: an auditory attention example. *Journal of Visualized Experiments*. 2012:e4262. [PubMed: 23128363]
- Lee AKC, Rajaram S, Xia J, Bharadwaj H, Larson E, Hämäläinen MS, Shinn-Cunningham BG. Auditory Selective Attention Reveals Preparatory Activity in Different Cortical Regions for Selection Based on Source Location and Source Pitch. *Frontiers in Neuroscience*. 2013; 6. [PubMed: 23386809]
- Logothetis NK. What we can do and what we cannot do with fMRI. *Nature*. 2008; 453:869–878. [PubMed: 18548064]
- Mathiak K, Rapp A, Kircher TTJ, Grodd W, Hertrich I, Weiskopf N, Lutzenberger W, Ackermann H. Mismatch responses to randomized gradient switching noise as reflected by fMRI and whole-head magnetoencephalography. *Human Brain Mapping*. 2002; 16:190–195. [PubMed: 12112773]
- Mayer AR, Franco AR, Harrington DL. Neuronal modulation of auditory attention by informative and uninformative spatial cues. *Human Brain Mapping*. 2009; 30:1652–1666. [PubMed: 18661505]
- Mayer AR, Harrington D, Adair JC, Lee R. The neural networks underlying endogenous auditory covert orienting and reorienting. *NeuroImage*. 2006; 30:938–949. [PubMed: 16388970]
- McLachlan N, Wilson S. The central role of recognition in auditory perception: A neurobiological model. *Psychological Review*. 2010; 117:175–196. [PubMed: 20063967]
- Mesgarani N, Chang EF. Selective cortical representation of attended speaker in multi-talker speech perception. *Nature*. 2012:1–5.

- Mishkin M, Ungerleider L, Macko K. Object Vision and Spatial Vision - 2 Cortical Pathways. *Trends in Neurosciences*. 1983; 6:414–417.
- Molholm S, Martinez A, Ritter W, Javitt DC, Foxe JJ. The neural circuitry of pre-attentive auditory change-detection: an fMRI study of pitch and duration mismatch negativity generators. *Cerebral Cortex*. 2005; 15:545–551. [PubMed: 15342438]
- Murray MM, Brunet D, Michel CM. Topographic ERP Analyses: A Step-by-Step Tutorial Review. *Brain Topography*. 2008; 20:249–264. [PubMed: 18347966]
- Näätänen R. The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral and Brain Science*. 1990; 13:201–88.
- Novitski N, Alho K, Korzyukov O, Carlson S, Martinkauppi S, Escera C, Rinne T, Aronen HJ, Näätänen R. Effects of acoustic gradient noise from functional magnetic resonance imaging on auditory processing as reflected by event-related brain potentials. *NeuroImage*. 2001; 14:244–251. [PubMed: 11525334]
- Paltoglou AE, Sumner CJ, Hall DA. Examining the role of frequency specificity in the enhancement and suppression of human cortical activity by auditory selective attention. *Hearing Research*. 2009; 257:106–118. [PubMed: 19706320]
- Paltoglou AE, Sumner CJ, Hall DA. Mapping feature-sensitivity and attentional modulation in human auditory cortex with functional magnetic resonance imaging. *European Journal of Neuroscience*. 2011; 33:1733–1741. [PubMed: 21447093]
- Peelle JE, Eason RJ, Schmitter S, Schwarzbauer C, Davis MH. Evaluating an acoustically quiet EPI sequence for use in fMRI studies of speech and auditory processing. *NeuroImage*. 2010; 52:1410–1419. [PubMed: 20483377]
- Petkov C, Kang X, Alho K, Bertrand O, Yund E, Woods D. Attentional modulation of human auditory cortex. *Nature Neuroscience*. 2004; 7:658–663.
- Posner MI. Orienting of attention. *Quarterly Journal of Experimental Psychology*. 1980; 32:3–25. [PubMed: 7367577]
- Power AJ, Foxe JJ, Forde EJ, Reilly RB, Lalor EC. At what time is the cocktail party? A late locus of selective attention to natural speech. *European Journal of Neuroscience*. 2012; 35:1497–1503. [PubMed: 22462504]
- Rauschecker JP. Cortical processing of complex sounds. *Current Opinion in Neurobiology*. 1998; 8:516–521. [PubMed: 9751652]
- Rauschecker JP, Scott SK. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nature Neuroscience*. 2009; 12:718–724.
- Rinne T. Activations of human auditory cortex during visual and auditory selective attention tasks with varying difficulty. *The Open Neuroimaging Journal*. 2010; 4:187–193. [PubMed: 21760872]
- Rinne T, Alho K, Ilmoniemi RJ, Virtanen J, Näätänen R. Separate time behaviors of the temporal and frontal mismatch negativity sources. *NeuroImage*. 2000; 12:14–19. [PubMed: 10875898]
- Rinne T, Balk MH, Koistinen S, Autti T, Alho K, Sams M. Auditory Selective Attention Modulates Activation of Human Inferior Colliculus. *Journal of Neurophysiology*. 2008; 100:3323–3327. [PubMed: 18922948]
- Rinne T, Pekkola J, Degerman A, Autti T, Jääskeläinen IP, Sams M, Alho K. Modulation of auditory cortex activation by sound presentation rate and attention. *Human Brain Mapping*. 2005; 26:94–99. [PubMed: 15852467]
- Roberts KL, Summerfield AQ, Hall DA. Presentation modality influences behavioral measures of alerting, orienting, and executive control. *Journal of the International Neuropsychological Society : JINS*. 2006; 12:485–492. [PubMed: 16981600]
- Roberts KL, Summerfield AQ, Hall DA. Covert auditory spatial orienting: An evaluation of the spatial relevance hypothesis. *Journal of Experimental Psychology-Human Perception and Performance*. 2009; 35:1178–1191. [PubMed: 19653757]
- Rosano C, Krisky CM, Welling JS, Eddy WF, Luna B, Thulborn KR, Sweeney JA. Pursuit and saccadic eye movement subregions in human frontal eye field: a high-resolution fMRI investigation. *Cerebral Cortex*. 2002; 12:107–115. [PubMed: 11739259]

- Ruytjens L, Willemsen A, Van Dijk P, Wit H, Albers F. Functional imaging of the central auditory system using PET. *Acta Oto-Laryngologica*. 2006; 126:1236–1244. [PubMed: 17101583]
- Salmi J, Rinne T, Degerman A, Alho K. Orienting and maintenance of spatial attention in audition and vision: an event-related brain potential study. *European Journal of Neuroscience*. 2007a; 25:3725–3733. [PubMed: 17610592]
- Salmi J, Rinne T, Degerman A, Salonen O, Alho K. Orienting and maintenance of spatial attention in audition and vision: multimodal and modality-specific brain activations. *Brain Structure and Function*. 2007b; 212:181–194. [PubMed: 17717689]
- Salmi J, Rinne T, Koistinen S, Salonen O, Alho K. Brain networks of bottom-up triggered and top-down controlled shifting of auditory attention. *Brain Research*. 2009; 1286:155–164. [PubMed: 19577551]
- Serences JT, Schwarzbach J, Courtney SM, Golay X, Yantis S. Control of object-based attention in human cortex. *Cerebral Cortex*. 2004; 14:1346–1357. [PubMed: 15166105]
- Sharon D, Hämäläinen MS, Tootell RBH, Halgren E, Belliveau JW. The advantage of combining MEG and EEG: comparison to fMRI in focally stimulated visual cortex. *NeuroImage*. 2007; 36:1225–1235. [PubMed: 17532230]
- Shinn-Cunningham BG. Object-based auditory and visual attention. *Trends in Cognitive Sciences*. 2008; 12:182–186. [PubMed: 18396091]
- Shomstein S, Yantis S. Parietal cortex mediates voluntary control of spatial and nonspatial auditory attention. *Journal of Neuroscience*. 2006; 26:435–439. [PubMed: 16407540]
- Slagter H, Weissman D, Giesbrecht B, Kenemans J, Mangun G, Kok A, Woldorff M. Brain regions activated by endogenous preparatory set shifting as revealed by fMRI. *Cognitive, Affective & Behavioral Neuroscience*. 2006; 6:175–189.
- Smith DV, Davis B, Niu K, Healy EW, Bonilha L, Fridriksson J, Morgan PS, Rorden C. Spatial attention evokes similar activation patterns for visual and auditory stimuli. *Journal of Cognitive Neuroscience*. 2010; 22:347–361. [PubMed: 19400684]
- Szczepanski SM, Kastner S. Shifting attentional priorities: control of spatial attention through hemispheric competition. *Journal of Neuroscience*. 2013; 33:5411–5421. [PubMed: 23516306]
- Szczepanski SM, Konen CS, Kastner S. Mechanisms of spatial attention control in frontal and parietal cortex. *Journal of Neuroscience*. 2010; 30:148–160. [PubMed: 20053897]
- Tark KJ, Curtis CE. Persistent neural activity in the human frontal cortex when maintaining space that is off the map. *Nature Neuroscience*. 2009; 12:1463–1468.
- Teki S, Chait M, Kumar S, von Kriegstein K, Griffiths TD. Brain bases for auditory stimulus-driven figure-ground segregation. *Journal of Neuroscience*. 2011; 31:164–171. [PubMed: 21209201]
- Urbach TP, Kutas M. Interpreting event-related brain potential (ERP) distributions: Implications of baseline potentials and variability with application to amplitude normalization by vector scaling. *Biological Psychology*. 2006; 72:333–343. [PubMed: 16446023]
- Uutela K, Hämäläinen M, Somersalo E. Visualization of Magnetoencephalographic Data Using Minimum Current Estimates. *NeuroImage*. 1999; 10:173–180. [PubMed: 10417249]
- Voisin J. Listening in Silence Activates Auditory Areas: A Functional Magnetic Resonance Imaging Study. *Journal of Neuroscience*. 2006; 26:273–278. [PubMed: 16399697]
- Wardak C, Ibos G, Duhamel JR, Olivier E. Contribution of the monkey frontal eye field to covert visual attention. *Journal of Neuroscience*. 2006; 26:4228–4235. [PubMed: 16624943]
- Warren J, Griffiths T. Distinct mechanisms for processing spatial sequences and pitch sequences in the human auditory brain. *Journal of Neuroscience*. 2003; 23:5799–5804. [PubMed: 12843284]
- Watkins S, Dalton P, Lavie N, Rees G. Brain mechanisms mediating auditory attentional capture in humans. *Cerebral Cortex*. 2007; 17:1694–1700. [PubMed: 16990437]
- Wilcox T, Stubbs J, Hirshkowitz A, Boas DA. Functional activation of the infant cortex during object processing. *NeuroImage*. 2012; 62:1833–1840. [PubMed: 22634218]
- Woldorff MG, Gallen CC, Hampson SA, Hillyard SA, Pantev C, Sobel D, Bloom FE. Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proceedings of the National Academy of Sciences of the United States of America*. 1993; 90:8722–8726. [PubMed: 8378354]



- Woods DL, Herron TJ, Cate AD, Yund EW, Stecker GC, Rinne T, Kang X. Functional Properties of Human Auditory Cortical Fields. *Frontiers in Systems Neuroscience*. 2010; 4
- Woods DL, Stecker GC, Rinne T, Herron TJ, Cate AD, Yund EW, Liao I, Kang X. Functional maps of human auditory cortex: effects of acoustic features and attention. *PLoS ONE*. 2009; 4:e5183. [PubMed: 19365552]
- Xiang J, Simon J, Elhilali M. Competing streams at the cocktail party: exploring the mechanisms of attention and temporal integration. *Journal of Neuroscience*. 2010; 30:12084–12093. [PubMed: 20826671]
- Yang Z, Mayer AR. An event-related fMRI study of exogenous orienting across vision and audition. *Human Brain Mapping*. 2013:n/a–n/a.
- Yantis S. The Neural Basis of Selective Attention: Cortical Sources and Targets of Attentional Modulation. *Current Directions in Psychological Science*. 2008; 17:86–90. [PubMed: 19444327]
- Yantis S, Schwarzbach J, Serences JT, Carlson RL, Steinmetz MA, Pekar JJ, Courtney SM. Transient neural activity in human parietal cortex during spatial attention shifts. *Nature Neuroscience*. 2002; 5:995–1002.
- Zion Golumbic EM, Cogan GB, Schroeder CE, Poeppel D. Visual Input Enhances Selective Speech Envelope Tracking in Auditory Cortex at a “Cocktail Party”. *Journal of Neuroscience*. 2013; 33:1417–1426. [PubMed: 23345218]

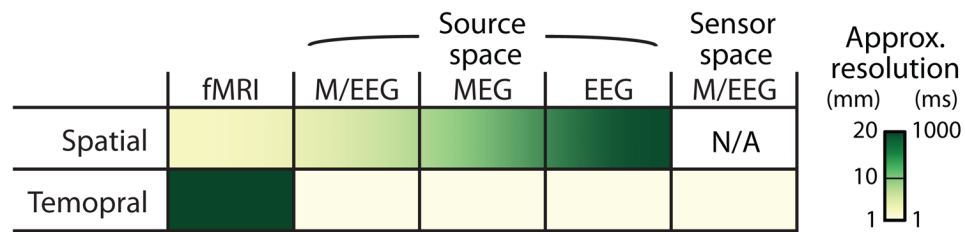
**Highlights**

Different neuroimaging techniques have different strengths and challenges.

Successful neuroimaging experiments require careful design.

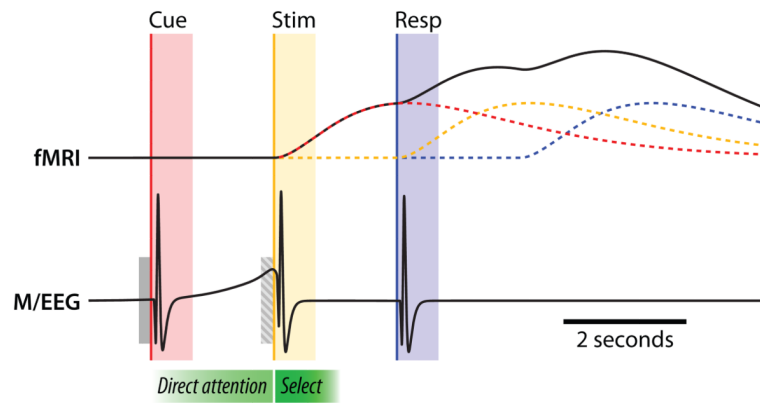
Attention modulates auditory sensory areas as well as higher cortical regions.

Auditory and visual attention share a similar supramodal attentional network.



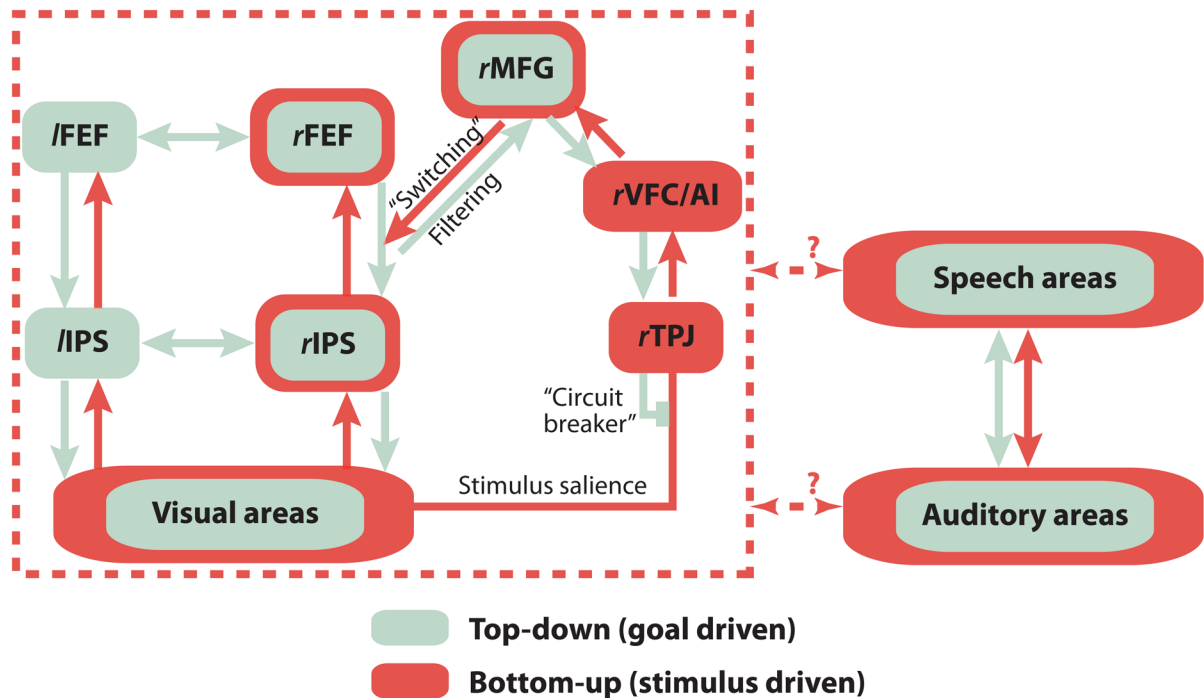
**Figure 1.**

Approximate spatial resolution and temporal resolution differ dramatically across imaging modalities. While fMRI has excellent spatial resolution (sub-centimeter) compared to M/EEG (around a centimeter), it has comparatively poor temporal resolution (seconds versus milliseconds, respectively). Sensor space analysis is based directly on the field topographical patterns (see Section 2.2.2), while source space analysis seeks to map the topographical patterns to the underlying neural sources analysis using ECD or inverse modeling (see Section 2.2.3).



**Figure 2.**

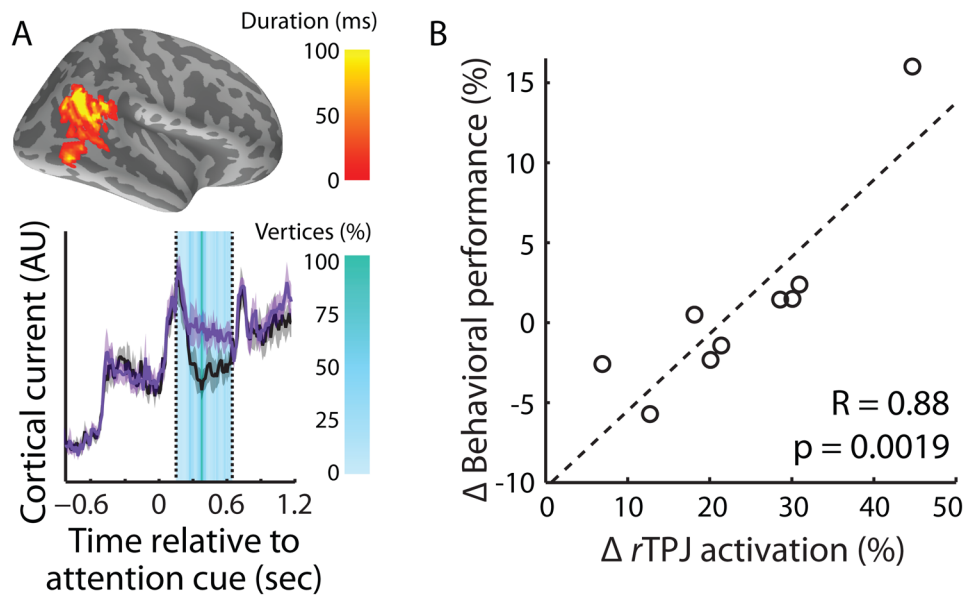
Schematic showing simplified, idealized signals that could be obtained using fMRI (BOLD response) and M/EEG (field strength or electric potentials) in a sample experiment. The three parts of a trial (cue, red; stimulus, yellow; and response, blue) all elicit a neural response. In the BOLD signal, these responses overlap and sum as a result of their extended time courses (note also the 2 s delay; Boynton et al., 1996), making timing-based analysis difficult. A catch trial (presenting a cue without a subsequent stimulus) can be used to isolate the response in the cue-stimulus interval [i.e., removing the stimulus (yellow) contribution in the overall (black) time course]. With M/EEG data, the responses are brief and separable, allowing identification of the preparatory attentional signal preceding the stimulus period (yellow). Gray bars denote the common practice of baseline correction using the pre-cue (solid) or the pre-stimulus (hatch) interval. See Section 3 for in-depth discussion.



**Figure 3.**

The top-down and bottom-up attention network involved in visual attention (modified from (Corbetta et al., 2008) likely operates supramodally. Evidence suggests that auditory tasks engage many of these same areas as visual tasks, although the connections between various nodes in the network remain unknown.





**Figure 4.**

In one M/EEG experiment, the right temporoparietal junction (rTPJ) was significantly more active when subjects switched attention (purple trace) compared to when they maintained attention (black trace; reproduced from Larson et al., 2013). The duration of each vertex in the cluster is shown (A, top) alongside the time evolution of the neural activity (colored by the percentage of vertices in the cluster was significant; A, bottom). Differential activation in rTPJ was also correlated with behavioral performance differences in the task (B). These types of neural-behavioral correlations help establish the role of various regions in switching auditory spatial attention.