2019

Cortical mechanisms underlying auditory spatial and non-spatial selective attention

BOSTON UNIVERSITY

COLLEGE OF ENGINEERING

Dissertation

CORTICAL MECHANISMS UNDERLYING AUDITORY SPATIAL AND NON-SPATIAL SELECTIVE ATTENTION

by

YUQI DENG

B.E., University of Hong Kong, 2014
M.S., Boston University, 2017

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2019

Approved by

First Reader
Barbara G. Shinn-Cunningham, Ph.D.
Research Professor of Biomedical Engineering
Boston University

Professor of Biomedical Engineering, Psychology, and Electrical and Computer Engineering
Carnegie Mellon University

Second Reader
H. Steven Colburn, Ph.D.
Professor of Biomedical Engineering
Professor of Electrical and Computer Engineering

Third Reader
Nancy Kopell, Ph.D.
William Fairfield Warren Distinguished Professor
Professor of Mathematics and Statistics
Professor of Biomedical Engineering
Professor of Pharmacology and Experimental Therapeutics

Fourth Reader
Inyong Choi, Ph.D.
Assistant Professor of Communication Sciences and Disorders
The University of Iowa

Fifth Reader
Robert M.G. Reinhart, Ph.D.
Assistant Professor of Psychological and Brain Sciences

ACKNOWLEDGMENTS

Foremost I would like to thank my advisor Dr. Barbara Shinn-Cunningham, who is not only a visionary scientist but also a caring mentor. I miss our weekly meetings, which are forever inspiring and cheerful. Many thanks to Barb for taking me under her wings, discovering my strength and potentials and offering me generous guidance. I would also like to express my gratitude to my committee members: Dr. Steve Colburn, Dr. Nancy Kopell, Dr. Inyong Choi and Dr. Robert Reinhart for their continuous support and insightful discussions.

I would like to express my gratitude for many brilliant scientists who have mentored me: Dr. Inyong Choi, Dr. Owen Brimijoin, Dr. Robert Baumgartner and Dr. Robert Reinhart. My special thanks goes to Inyong, who patiently taught me the analysis skills when I was a beginner, and continued to be a motivating and helpful mentor throughout my study. I am also grateful to have worked with smart and fun lab members. This experience wouldn’t have been colorful without them.

Lastly, I would like to thank my family, who trust me with all my endeavors and supported me through my ups and downs.

CORTICAL MECHANISMS UNDERLYING AUDITORY SPATIAL AND NON-SPATIAL SELECTIVE ATTENTION

YUQI DENG

Boston University College of Engineering, 2019

Major Professor: Barbara G. Shinn-Cunningham, Ph.D., Research Professor of Biomedical Engineering, Boston University; Professor of Biomedical Engineering, Psychology, and Electrical and Computer Engineering, Carnegie Mellon University

ABSTRACT

Despite the importance of auditory selective attention in everyday communication, the cortical mechanisms underlying the amazing ability of human brain to focus on a sound stimulus and suppress others are not well understood. Recent studies have led to the hypothesis that alpha band oscillation (8–14Hz) is a neural signature of multimodal spatial attention. Research in multiple sensory domains has shown that alpha synchronizes in the hemisphere contralateral to unattended stimuli and desynchronizes on the hemisphere contralateral to attended stimuli, suggesting it is a marker of an inhibition process for filtering out unattended stimuli. However, further research is needed to understand the possible functional role of these alpha oscillations as well as their correlation with other cortical activity. Moreover, it is not clear whether different forms of auditory attention employ different cortical mechanisms, mediated through different brain networks.

This study aims to combine brain stimulation methods (transcranial Direct/Alternative Current Stimulation) with electrophysiological measurements of electroencephalography (EEG) to measure and interpret the underlying cortical activity during different forms of auditory selective attention. More specifically, there are four studies, each of which employs behavioral tasks to test specific hypotheses. First, we studied alpha oscillatory activity during auditory spatial attention. Second, we compared and contrast cortical activity during auditory spatial and non-spatial attention. Third, we used brain stimulation to see if we can show a causal relationship between alpha oscillation and selective auditory attention performance. Lastly, we applied the existing results on alpha power to use it as a quantitative biomarker to indicate the level of spatial attention network engagement. Our results contributed to the growing body of knowledge about how the brain employs auditory selective attention for effective communication.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

EEG	Electroencephalography
ERP	Event Related Potential
fMRI	Functional Magnetic Resonance Imaging
HD - tACS	High Definition Transcranial Alternating Current Stimulation
HRTF	Head Related Transfer Function
IPS	Intra-parietal Sulcus
MEG	Magnetoencephalography
SRM	Spatial Release from Masking
tACS	Transcranial Alternating Current Stimulation
tDCS	Transcranial Direct Current Stimulation
tES	Transcranial Electrical Stimulation

INTRODUCTION

In an everyday social setting where the sound scene is a complex mixture of multiple conversations and background noises, auditory selective attention plays a crucial role for effective communication. It allows listeners to not only filter out unwanted sound sources but also dynamically shifts attention focus across different sounds based on their physical characteristics. Historically presented as the cocktail party problem (Cherry, 1953), auditory selective attention has attracted many research efforts to understand the underlying cognitive mechanisms. Previous studies have shined lights on this matter by investigating auditory selective attention focused upon different sound attributes, including spatial and non-spatial acoustic features (e.g. Mondor and Bregman, 1994; Kidd et al., 2005; Wu et al., 2007; Maddox and Shinn-Cunningham, 2012; Larson and Lee, 2014). However, despite the amount of interest, exactly how the brain manages to achieve auditory selective attention is still unclear.

(Shinn-Cunningham, 2008a) built a framework to understand auditory selective attention by extending theories from visual attention to the auditory domain. This method is supported by many literature findings that auditory and visual attention shares similar control network (e.g. Ward, 1994; Serences et al., 2004; Roberts and Hall, 2008; Luo et al., 2010; Green et al., 2011). Specifically, an influential hypothesis of attention proposes that the same spatial attention network in the parietal cortex (Farah et al., 1989), specifically in the intraparietal sulcus (IPS) (Weisz et al., 2013), controls spatial attention across multiple sensory modalities including vision and audition (Green et al., 2011; Salmi et al., 2009). Since visual attention has been more extensively studied than many other sensory modalities, current knowledge on vision provides important implications for auditory attention mechanisms.

A known phenomenon during visual spatial attention is parietal-occipital alpha lateralization specific to attention focus. Alpha oscillation is a brain oscillation at 8-14Hz that is hypothesized to be neural signature of an inhibition mechanism (Klimesch, 2012a). A number of visuospatial studies have shown that alpha oscillation synchronizes (increases in power compared to baseline) in the hemisphere ipsilateral to an attended stimulus and desynchronizes (decreases in power) contralateral to the attended side (Kelly, 2006a; van Diepen et al., 2016; Worden et al., 2000a) (see Figure 1 for typical alpha lateralization pattern). Given that parietal cortex generally represents contralateral exocentric space, this pattern is consistent with increased suppression of objects in the hemifield away from an attended object and decreased suppression of objects in the attended hemifield. Alpha lateralization has recently been observed in auditory spatial attention too (S. Banerjee et al., 2011; Mehraei et al., 2018; Wöstmann et al., 2016a). However, with limited number of studies available, many questions regarding the exact role alpha oscillation plays during auditory attention remains unanswered.

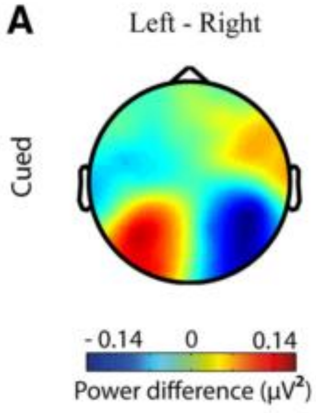

Figure 1. Adapted from (van Diepen et al., 2016). Typical alpha lateralization topography during visual spatial attention.

The aim of this thesis study is to further the understanding of the functions of alpha oscillation during different forms of auditory attention. Specifically, we investigated in correlative observations of alpha activities during auditory attention as well as causal evidence that alpha plays an active functional role during auditory attention (Figure 2). Electroencephalography (EEG) was used to observe correlative alpha activities during auditory attentional task due to the high temporal resolution and relatively low budget to operate. High-definition transcranial alternating current stimulation (HD-tACS) was used to investigate the causal relationship between alpha oscillation and attentional performance. HD-tACS was selected due to minimal invasiveness and high spatial resolution.

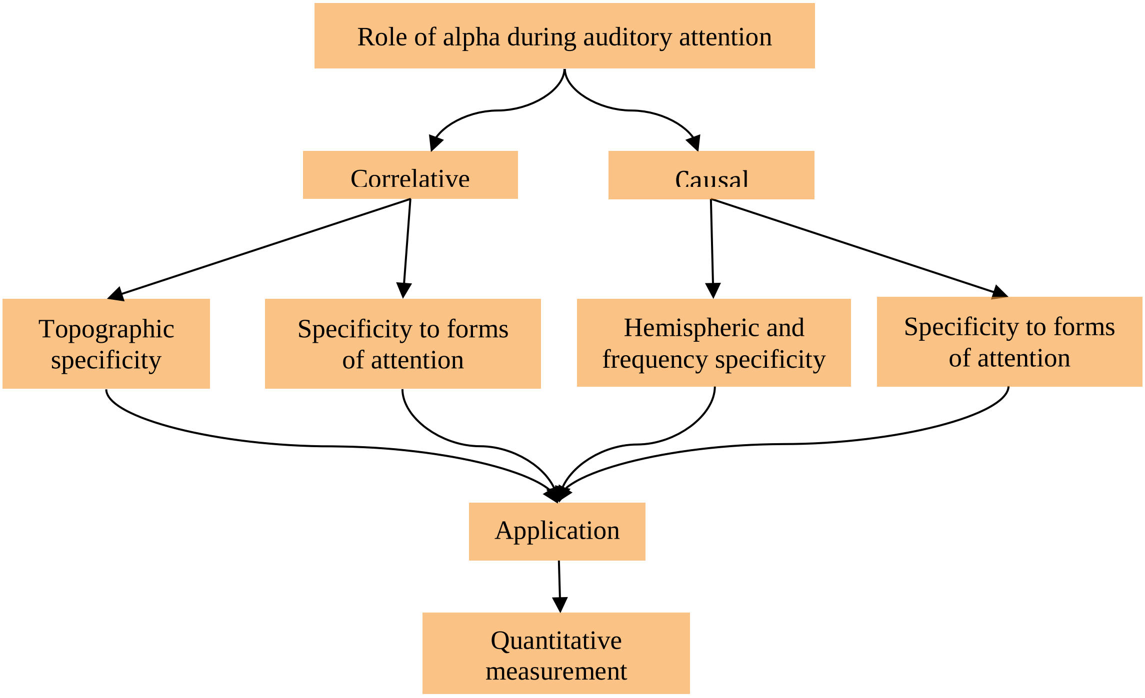

Figure 2. *Illustration of the structure of this thesis study. The role of alpha oscillation is explored from the correlational and causal aspect, with two specific*

questions answered within each aspect. Some potential applications of these findings are also studied and discussed.

In the correlative aspect, we aimed to study the alpha activities observed during auditory attention. In Chapter Two, we present a study conducted to investigate into topographic specificity of alpha oscillation to auditory spatial attention focus. While previous visuospatial attention studies have demonstrated that alpha lateralization pattern changes topographically when attention focus shifts in space (Samaha et al., 2016a; Worden et al., 2000a) (see Figure 3), no previous studies have shown that alpha oscillations induced by auditory attention lateralize topographically and reflect spatial attention focus. We hypothesize, based on the supramodal theory of the spatial attention network, that alpha will behave similarly during audiospatial attention and visuospatial attention. In addition, while inter-subject alpha peak frequency variability can be largely explained by inheritance (Bodenmann et al., 2009), intra-subject variability is not as well understood. It has been suggested that intra-subject alpha peak variability indicates different alpha generating networks (Haegens et al., 2014). Therefore, we will further explore whether multiple alpha generators exist during auditory spatial attention.

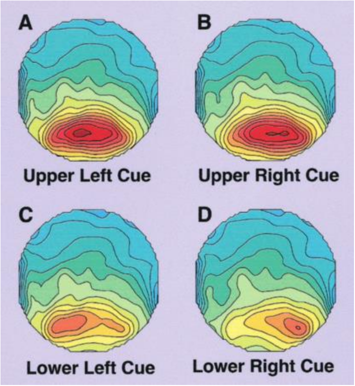

Figure 3. Adapted from (Worden. et al., 2000). Alpha band activity changes topographically when visuospatial attention focus shifts in space.

In addition, we aimed to study whether alpha activities are modulated differentially when auditory attention is based on different physical characteristics. In Chapter Two, we present a study investigating the differential cortical dynamics during auditory spatial and non-spatial attention. The role alpha oscillation plays during non-spatial attention is poorly understood compared to its role in spatial attention. Past work suggests that when continuous visuospatial attention is focused, non-spatial visual attention does not change alpha lateralization in the parietal regions (H A Slagter et al., 2016). However, some studies have claimed that alpha oscillation can be modulated by feature-based visual attention (John J. Foxe and Snyder, 2011). Therefore, it remains unclear whether non-spatial auditory attention can cause similar alpha lateralization modulation as spatial attention. Previous studies suggest that spatial and non-spatial attention engage completely different cortical networks (Lee et al., 2013). Based on this, we hypothesize that spatial and non-spatial auditory attention employs different neural mechanisms and behavioral strategies, which may be reflected in different patterns of induced and evoked cortical activity.

Besides correlative studies, we aimed to investigate in whether alpha oscillation plays causal functional roles during auditory attention. Advances in the use of non-invasive brain stimulation technologies open up the possibility of safely targeting certain cortical regions and causally changing electrophysiological responses in these regions. Transcranial direct and alternative current stimulation (tDCS & tACS) are non-invasive brain stimulation methods that have been used to introduce direct and alternating currents through the scalp to either activate or deactivate certain cortical regions (Reinhart et al., 2017). Previous studies using transcranial magnetic stimulation (TMS) have shown that disruption of alpha oscillation in the parietal region impairs behavioral performance during visuospatial selective attention (V Romei et al., 2010). Therefore, in Chapter Three, we present a study using brain stimulation method (HD-tACS) to improve understanding of how parietal cortices contribute to auditory spatial and non-spatial attention, and to demonstrate causal effects of parietal and temporal responses on neural signatures and behavior performance. Specifically, we hypothesize that the stimulation effect is hemispheric specific, where stimulating one hemisphere would disturb attention focuses on the contra-lateral hemifield to the stimulated hemisphere. We included a control group stimulated by theta frequency (6Hz) to show that the stimulation effect is frequency specific. Moreover, we included in this study both spatial attention tasks and non-spatial attention task to test the hypothesis that stimulating the parietal “where” pathway will influence auditory spatial attention, and that stimulating the temporal “what” pathway will influence auditory non-spatial attention.

With better understanding of alpha as a neural signature underlying auditory attention, question remains as to how these results could potentially benefits real-world applications. In visual spatial attention, the topographic specificity of alpha has been employed for development of brain computer interfaces (Samaha et al., 2016b). There has also been attempts to use alpha as a biomarker for clinical diagnosis (Koehler et al., 2009; Mathewson et al., 2012). Here, in Chapter Four, we used alpha as a quantitative measurement for the level of engagement of auditory spatial attention network. Together with event relation potentials (ERP) as another quantitative measurement, we tested the hypothesis that impoverished auditory spatial cues do not fully engage auditory spatial attention network as much as natural spatial cues.

Together, we hope results from this thesis study would provide important knowledge on the cortical activities underlying auditory attention based on various physical characteristics. Specifically, it was demonstrated both correlatively and causally that alpha oscillation plays differential functional roles during different forms of auditory attention.

CHAPTER ONE

Topographic Specificity of Alpha Oscillation to Auditory Spatial Attention Focuses

Yuqi Deng[1, 2], Inyong Choi[3], Barbara Shinn-Cunningham[1, 2]

Statement of permission:

This manuscript has been prepared for submission to a peer-reviewed journal.

[1] Department of Biomedical Engineering, Boston University, Boston, MA, USA, 02215; vydeng@bu.edu

[2] Center for Research in Sensory Communication and Emerging Neural Technology, Boston University, 610 Commonwealth Ave, Boston, MA, USA, 02215; shinn@bu.edu

[3] Department of Communication Sciences and Disorders, University of Iowa, Iowa City, IA, 52242; inyong-choi@uiowa.edu

Abstract

Visual and somatosensory spatial attention has been reported to induce alpha oscillation (7–14Hz) lateralization that is topographically specific to attention focuses. However, the topographic specificity of alpha power in the auditory modality is less clear. Moreover, most previous studies on alpha oscillation used a fixed spectral band uniformly (usually 7–14Hz) to extract alpha power without examining the variability of alpha peak frequencies across cortical regions. Here, we recorded electroencephalography (EEG) in human subjects while they were performing a spatial attention task, in which they were cued to attend to one of five spatial locations along the azimuth. Besides inter-individual differences, during the attentional preparation period we found intra-individual differences in peak alpha frequency between parietal and central cortex, suggesting multiple generating sources. We also examined hemispheric lateralization of individualized alpha power and found alpha topographies specific to attention focuses. While alpha synchronization and desynchronization was observed in both hemispheres, more attentional modulation was observed in the left hemisphere when attention focuses shifted. Our results on topographic specificity and hemispheric asymmetry of alpha modulation supports the theory of a right hemisphere dominated supramodal attention network controlling spatial attention in multiple sensory modalities.

Keywords

EEG, auditory, spatial, attention, alpha, oscillation, asymmetry

1 Introduction

Auditory spatial attention enables listeners to process auditory information only from the relevant location and thus is crucial for effective communication in complex auditory scenes. However, although the underlying mechanisms of spatial attention in the visual modality has been extensively studied, that of the spatial attention in auditory modality is less well understood.

In the visual modality, it is established that top down modulation of spatial attention is controlled by the fronto-parietal attention network (Buschman and Miller, 2007; Corbetta and Shulman, 2002; He et al., 2007). An influential hypothesis of spatial attention proposes that the same fronto-parietal attention network controls spatial attention across all sensory modalities (Brunetti et al., 2008; Farah et al., 1989; Wu et al., 2007). There are both behavioral and physiological evidence supporting this theory (Macaluso et al., 2000; Pavani et al., 2003; Wu et al., 2007). Recently, new evidence has been emerging showing that spatial attention in multiple sensory domains including auditory can modulate parietal-occipital alpha oscillation in similar ways (S. Banerjee et al., 2011), and the hypothesized generator of the alpha oscillation is the intraparietal sulcus (IPS) that is part of the fronto-parietal attention network (Capotosto et al., 2009; Weisz and Obleser, 2014).

Alpha oscillation is a brain oscillation at 7–14Hz that was historically thought to be signature of the idling state. However, more recently it has been hypothesized to be a neural signature of an inhibition mechanism (John J Foxe and Snyder, 2011; Klimesch et al., 2007a; V. Romei et al., 2010). A number of studies have shown that alpha oscillation synchronizes (increases in power compared to baseline) in the hemisphere ipsilateral to an attended stimulus and desynchronizes (decreases in power compared to baseline) contralateral to the attended side in visuospatial attention (Kelly, 2006a; Thut et al., 2006; Worden et al., 2000b; Wöstmann et al., 2016a) and audiospatial attention (Frey et al., 2014; Strauß et al., 2014). Given that parietal cortex generally represents contralateral exocentric space, this pattern is consistent with increased suppression of objects in the hemifield opposite to an attended object and decreased suppression of objects in the attended hemifield (Huang et al., 2014).

Although alpha lateralization has recently been observed in audiospatial attention too, details about its modulation mechanisms remain unclear. While previous visuospatial attention studies have shown that alpha lateralization pattern changes topographically with different attention focuses in space (Rihs et al., 2007a; Samaha et al., 2016b; van Gerven and Jensen, 2009; Worden et al., 2000b), none have proved that alpha oscillations induced by auditory attention lateralize topographically and reflect spatial attention focus. According to the theory that alpha oscillation is an active control of spatial attention by a supramodal network, we hypothesized that alpha power in the parietal-occipital region changes topographically to attention focus space similarly to that reported in visuospatial attention.

It is established that the spatial attention network is lateralized to be right hemisphere dominant (De Schotten et al., 2011; Shulman et al., 2010). A “Neglect model” has proposed that right parietal cortex controls auditory attention to both hemifields, while left parietal cortex only controls right hemisfield (Teshiba et al., 2013) with support from previous neuroimaging studies (Kaiser et al., 2000; Schonwiesner et al., 2006). The model explains why hemineglect happens more often with lesion in right parietal cortex than left parietal cortex. Based on the neglect model, Huang et al hypothesized that the left IPS regulate auditory information from only the opposite (right) hemifield while the right IPS would regulation spatial attention to both hemifields (Huang et al., 2014; Mesulam, 1999; Okazaki et al., 2015; Pouget and Driver, 2000). Their hypothesis was supported by comparing contralateral and ipsilateral phase locking activities of left and right IPS to auditory cortex during spatial attention.

However, past studies that have investigated into the hemispherical asymmetry of alpha oscillation show controversial results. While a visuospatial attention study and a tactile spatial attention study reveal that left hemisphere invoked more alpha modulation than right hemisphere (Saskia Haegens et al., 2011; Ikkai et al., 2016a), another visuospatial attention study reveal the opposite (Kelly, 2006a). Moreover, the typical right hemisphere dominance in spatial attention does not hold true for all individuals. Handedness has been shown to be related to different hemispheric asymmetries. The probability of atypical hemispheric dominance is strongly associated with left-handedness (Cai et al., 2013; Knecht, 2000). In spite of the possible influence of handedness on spatial attention asymmetry, left-handed subjects are underrepresented in published brain oscillation studies. Most previous studies on alpha oscillation either recruited only right-handed subjects or did not specify handedness preference of subjects (Willems et al., 2014).

In this study, we aim to further explore the hemispheric asymmetry associated with alpha modulation. According to the neglect model and the assumption that IPS is the primary alpha generator, we hypothesize in this study that alpha power during spatial attention is asymmetric between left and right hemisphere. Our specific hypothesis is that when attention focus shifts from left to right alpha power on the right hemisphere will show little variation since both hemifields are represented by it, whereas the left hemisphere alpha will decrease in power. In addition, in this study we recruited both left- and right-handed individuals and further investigated into the individual differences of alpha oscillation. In addition, despite the assumption that alpha oscillation generated by the IPS is a neural signature of spatial attention, there may exist multiple generating sources of alpha oscillation. It has been suggested that intra-subject alpha peak variability indicates different alpha generating networks (Haegens et al., 2014). Thus, another question we ask is whether intra-subject variability of alpha peak frequency exist during auditory spatial attention.

To test our hypothesis, we designed an auditory attention task that cues the listeners to actively attend to one of the five spatial locations along the azimuth. While listeners were actively engaged in the task we measured their behavioral performance along with the electrophysiological signals by electroencephalography (EEG). We investigated into inter- and intra- individual differences in alpha peak frequency, alpha lateralization topography as well as hemispheric asymmetry in alpha modulation.

2 Materials and Methods

2.1 Participants

Thirty subjects (14 females, 18-30 years of age) participated in this study. All subjects had normal hearing (hearing thresholds better than 20dB at pure tone frequencies between 250Hz and 8 kHz). All gave informed consent as approved by the Boston University Institutional Review Board. Two subjects were excluded from the study due to inability to perform the task (percentage of correct response equal to chance level). Subjects were asked to fill out an Edinburgh handedness inventory questionnaire (Oldfield, 1971) to determine their handedness preference. 14 out of the 28 remaining subjects were right-handed while the other 14 subjects are left-handed.

2.2 Paradigm

Participants performed a spatial attention task (Figure 4). They were cued by an auditory cue to listen to one of five locations in space. The sound stimuli started 2s after the cue. The sound stimuli consisted of a target sound from the cued location and a distractor sound from a random location that differed from that of the target sound. Subjects were asked to report the target sound sequence.

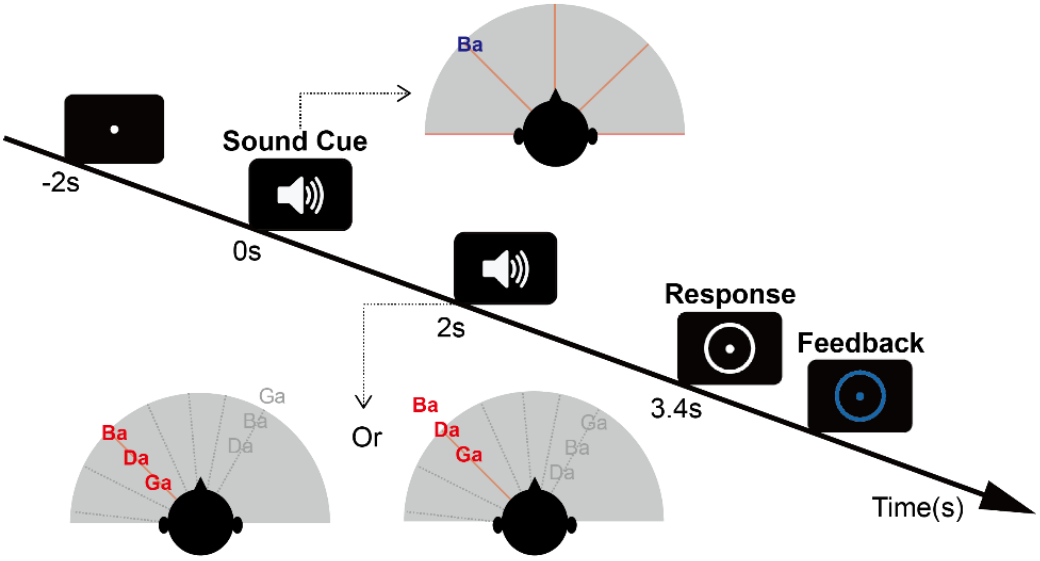

Figure 4. Trial design. A fixation dot appears at the center of the screen to instruct the listeners to fixate their gaze. An auditory cue (colored blue) of 400ms duration is presented 2 s later from one of five locations in space to indicate the location to be attended. After attentional preparation (0 s – 2 s), a target sound (colored red) made up of random sequences of /ba/, /da/ & /ga/ appears from the cued direction while a distractor stream (colored grey) appears from a different location than the target. Solid lines indicated the five possible direction of target; dashed lines indicated the possible directions of the distractor. The target and distractor are separated by 200ms. In half the trials the target leads the Distractor by 200ms, in the other half of the trials the Distractor leads. Listeners type in the sequence of the target after all streams finished playing and a white circle appears on the screen (at 3.4 s). Immediately after the response is given, a feedback of either blue or red circle appears to inform the correctness of response.

The target and distractor sound are separated by 200ms. To make sure that listeners engaged spatial attention rather than temporal attention, the onset times of the syllables in the target and distractor streams were pseudo-randomized so that in half of the trials, the target stream onset leads the distractor by 200ms, and in the other half of the trials the distractor stream leads. Subjects received feedback at the end of each trial about whether or not they correctly reported the syllables in the target stream. Participants perform 12 blocks of this task. Each block consists of 40 trials. In total each participant performed 480 trials, with each one of the five attended locations consisting of 96 trials. The stimulus sound level was at 70 dB SPL.

Auditory cue stimulus was a single syllable (/ba/). Target and distractor stimulus were three-syllable sequences made up of syllables (/ba/, /da/, & /ga/) spoken by the same female talker. The spatial effects of the stimuli was simulated by manipulating the interaural time difference (ITD) (Wightman and Kistler, 1992). Target sequences could be from 1 of 5 locations in space (ITDs: -600, -250, 0, 250, 600µs). The corresponding angular locations were estimated to be (-60°, -25°, 0°, 25°, 60°) on the azimuth with zero elevation (Smith and Price, 2014). The distractor stream was spatially separated from the target by one of 8 possible separations (ITDs: -600, -450, -300, -150, 150, 300, 450, 600µs). Only those distractor locations that were in the frontal -90 to 90 azimuth were included (-600µs to 600µs), regardless of where the target was. For both the target and distractor streams, syllables were separated by 500ms, onset to onset. The syllables of the target stream and the syllables of the distractor stream were temporally interdigitated by 200ms.

2.3 EEG data acquisition and preprocessing

EEG data were recorded with 64-channel Biosemi ActiveTwo system in an Eckel sound proof booth while participants were performing the tasks. Two additional reference electrodes were placed on the mastoids. The stimulus timing was controlled by Matlab (MathWorks, Natick, MA) with Psychtoolbox 3 extension (Brainard, 1997). A part of EEG analyses (e.g., plotting scalp topographies) was performed with EEGlab toolbox (Delorme and Makeig, 2004a) and Fieldtrip toolbox (Oostenveld et al., 2011). EEG data from the correct trials were referenced against the average of the mastoid channels and down-sampled to 256Hz. EEG data was then epoched according to the sound cue onset and then baseline corrected. The baseline period was chosen to be 100ms before cue onset. Trials were baselined corrected by subtracting the mean from the baseline period and then divided by the standard deviation of the baseline period. After baseline correction trials with a maximum absolute value over 80 microvolts were rejected (Delorme et al., 2007). Two subjects with excessive artifacts are removed from further EEG analysis (less than 60% trials remaining in at least one condition after artifact rejection). For the rest of the 26 subjects, after artifact rejection there are at least 92 trials remaining for each condition. To equate the number of trials, 92 trials were selected within each condition for each subject.

2.4 Data analysis

2.4.1 Behavioral data analysis:

We calculated the percentage of correctly recalled syllable for each one of the three syllables in the target stream. The conditions include 5 possible target locations, 3 sequence positions of the syllable and whether the target or the distractor stream is leading. Conditions with different relative locations of the distractor are collapsed for the statistical analysis. We conducted an ANOVA analysis on the percentage of correct response followed by post-hoc t-tests. Due to limited sample size (N=28), the normality of samples was checked by Lilliefors normality test (p>0.05) before performing parametric statistical tests. To correct for multiple comparisons, Benjamini-Hochberg FDR correction was performed for post-hoc tests.

2.4.2 Event related potential (ERP) analysis

For event related potential (ERP) analysis, the data from channel Cz is bandpass filtered 1–20Hz and baselined corrected against 100ms before the sound stimuli onset for each syllable. The local minimum value within the window of 50–150ms after the stimulus onset is taken as the ERP N1 amplitude.

2.4.3 Spectral analysis for inter and intra-individual differences:

Power spectra were calculated on 1s segment of each trial before the stimulus onset to avoid including the evoked activities. The data segments were zero padded to achieve a resolution of 0.1Hz. Trials within each condition was averaged (96 trials per condition) to get spectrum data for each EEG channel. Electrodes were grouped into the frontal group, fronto-central group and parieto-occipital group based on their locations on cortical surfaces (Figure 4A). Then the resulting power spectra were averaged within each electrode group to improve the signal-to-noise ratio. The peak alpha frequency for each condition was determined by finding the local maxima of the power spectra within the 7–14Hz band. The minimum peak width that can be detected was set at 0.6Hz to exclude potentially spurious peaks from noise. If there were multiple peaks in the alpha range, the peak with the maximum height was selected. Due to the lack of alpha peak detection in the frontal group for more than half of the subjects, alpha from the frontal group was excluded from further analysis. If alpha peak cannot be detected in at least one of the conditions for either fronto-central or parieto-occipital group, the subject was excluded from spectral analysis. One subject from the left-handed group was excluded from spectral analysis due to a lack of peak detection in the 7–14Hz range based on the aforementioned peak detection criterion. Individual alpha frequency (IAF) was selected as the peak frequency detected in the average spectrum across all trials in the parietal channels. The IAF for each subject was calculated and later used for later alpha oscillation analysis.

2.4.4 Alpha oscillation analysis

The data is bandpass filtered to the alpha range (IAF ± 1Hz). Then a Hilbert transform is applied to extract the alpha energy envelope, and the magnitude of the transformed data is taken as the alpha power. Alpha power was calculated for each trial then averaged to get the induced alpha power (Snyder and Large, 2005). Alpha power for each electrode was then baseline corrected against 1s before cue onset and then normalized to yield spatial z-scores. Spatial z-score of alpha power for each electrode was calculated by subtracting the mean alpha power of all electrodes and normalized against the global field power (Murray et al., 2008; Skrandies, 1990). The time courses of alpha power spatial z-scores were calculated for each attention location.

2.4.5 Spatial clustering analysis

The group-level analysis of GFP normalized alpha power was performed with MATLAB using the FieldTrip toolbox. The alpha power time course of each subject for each attended location was averaged across the time points of the whole trial (0–3.4s). We compared the conditions when subjects are attending to the far left (-60°) and the far right (60°) and clustered the EEG channels based on channel proximity. A Monte Carlo permutation test with 1000 random iterations was performed to control for multiple comparisons. The statistics implemented in Fieldtrip for this purpose was dependent samples t statistics. The cluster-based control has a type I error level of $\alpha = 0.05$. Clusters of neighboring electrodes with alpha power changing significantly with attended locations were revealed by this analysis. The resulted clusters were then used as regions of interest (ROI) for further analysis.

2.4.6 Linear regression model

Alpha power within each cluster was averaged to extract the average time course within each cluster. The time courses of normalized alpha power were averaged every time window of 200ms throughout the whole trial. A linear regression analysis was performed for each time interval of the alpha power time courses on a group level. The independent variable was the attended location and the dependent variable is the magnitude of alpha power. ANOVA was applied to the linear regression model and Bonferroni correction was performed to control for multiple comparisons. This analysis resulted in time windows exhibiting significant variations in attended locations.

3 Results

3.1 Behavioral results

Successful allocation of attention was confirmed by accurate reporting of target from the cued direction and ignoring the distractor. All subjects were able to perform significantly above chance level (33%). Percentage of correct response was affected by syllable sequences (ANOVA; $F_{(2,839)}=11.1$, $p<0.001$) and the interaction between syllable sequences and leading stream identity (ANOVA; sequence × leading stream: $F_{(2,839)}=63.9$, $p<0.001$). ANOVA did not show effects by target locations ($F_{(4,839)}=1.14$, $p=0.34$) or the interaction between target locations with other factors (location × leading stream: $F_{(4,839)}=1.22$, $p=0.30$; location × sequence: $F_{(8,839)}=0.22$, $p=0.99$), indicating similar task load at different target locations.

We further did post-hoc test to investigate in details the interaction between target locations and leading stream. Multiple comparison was controlled by the Benjamini-Hochberg procedure. Figure 5 shows the percentage of correctly recalling each syllable in the target leading and distractor leading conditions. Paired comparison between each two syllables show that percentage of correct response decreases with syllable number in the condition when the leading stream is the target stream (comparing: syllable 1&2, $t_{(27)}=9.83$, $p<0.001$; syllable 1&3, $t_{(27)}=9.49$, $p<0.001$; syllable 2&3, $t_{(27)}=3.78$, $p<0.001$; paired t-test). The trend is reversed when the leading stream is the distractor stream (comparing: syllable 1&2, $t_{(27)}=4.94$, $p<0.001$; syllable 1&3, $t_{(27)}=6.99$, $p<0.001$; syllable 2&3, $t_{(27)}=3.80$, $p<0.001$; t-test).

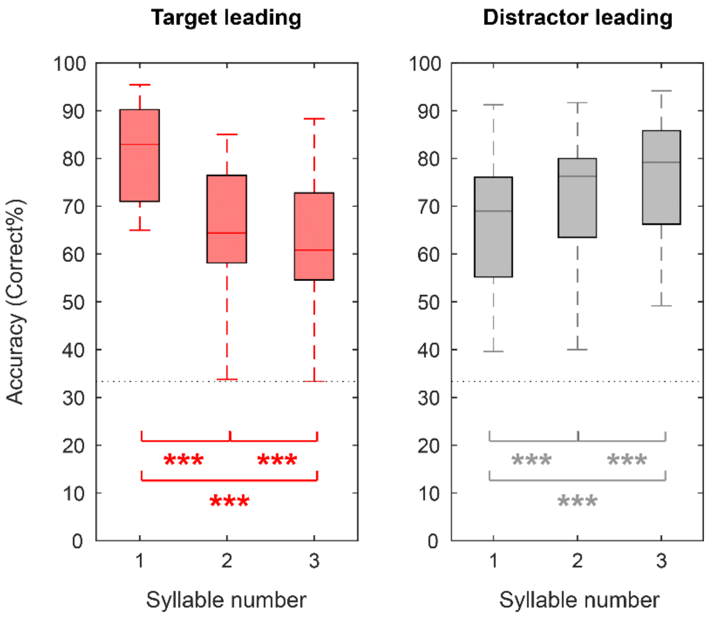

Figure 5. Behavioral performance of all subjects in target leading and distractor leading conditions. Label of x axis represents the percentage of correct response for each syllable of the target stream. Black dashed line represents chance performance (33%). Asterisks indicates statistical significance (p<0.05; ** p<0.01; *** p<0.001).*

To examine the effect of handedness on behavioral performance we performed a two sample t-test on the group level. The t-test did not reveal any difference between handedness groups ($t_{(14)}$=0.53, p=0.60).

3.2 Evoked event related potential (ERP) results

N1 peak magnitude of ERP to the auditory cue onset (0s) and target onset (2s) was calculated. The difference between these two conditions is that prior to the cue onset subjects were not aware of the location of the cue, whilst prior to the target onset subjects were already actively attending to the target location. Figure 6 shows the N1 amplitudes for each attentional condition. N1 amplitudes to the cue onset are generally larger than that to the target onset in all attention conditions. The amplitudes of N1 response to spatially lateralized sound stimulus are larger than the N1 amplitude response to the center (0°) for both the cue (-60°: $t_{(25)}=5.70$, $p<0.001$; 60°: $t_{(25)}=3.47$, $p=0.002$) and the target (-60°: $t_{(25)}=2.97$, $p=0.007$; 60°: $t_{(25)}=2.68$, $p=0.013$). This effect is less obvious when the sound stimuli is less lateralized along azimuth. Compared to midline, the N1 amplitudes at ±25° lateralization are trending statistically significant but not robust for both the cue (-25°: $t_{(25)}=1.95$, $p=0.063$; 25°: $t_{(25)}=2.63$, $p=0.014$) and the target (-25°: $t_{(25)}=0.82$, $p=0.42$; 25°: $t_{(25)}=2.16$, $p=0.04$).

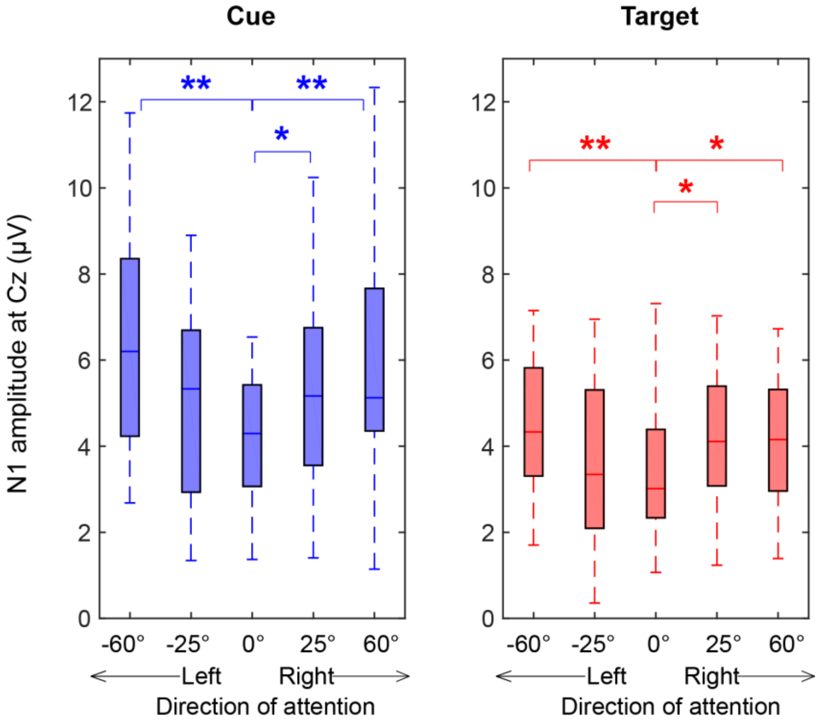

Figure 6. Amplitudes of N1 responses to auditory cue and target stimuli for each spatial attention location. Asterisks indicates statistical significance (p<0.05; ** p<0.01; *** p<0.001).*

3.3 Inter- and intra-individual differences in alpha peak frequency

Spectral analysis rendered peak frequency in the alpha range for each trial at each region of interest (ROI): the fronto-central ROI and the parieto-occipital ROI. Mean central-frontal alpha peak frequency of all subjects is 10.08Hz with a standard deviation of 0.6Hz across subjects while the mean at parietal-occipital region is 10.19Hz with a standard deviation of 0.55Hz across subjects. Alpha activities was observed in the frontal region in some trials but was not consistent.

We found that target locations do not have a significant effect on alpha peak frequency variation ($F_{(4,249)}$=1.63, p=0.12), but that subject identity ($F_{(24,249)}$=220, p<0.001) and ROI ($F_{(1,249)}$=59.27, p<0.001) do have significant main effects. Moreover, the interactions between subject identity and ROI ($F_{(24,249)}$==6.61, p<0.001) and the interaction between subject identity and target locations ($F_{(96,249)}$=1.51, p=0.022) are also significant. The interaction between target location and ROI is not significant ($F_{(4,249)}$=0.80, p=0.53). These results suggest that alpha peak frequencies vary across subjects, while different cortical regions vary in their peak alpha frequencies within subject. Post-hoc t-test reveals that the alpha peak frequency in the parieto-occipital region is larger than that in the fronto-central region ($t_{(24)}$=2.99, p=0.006; paired t-test) (Figure 7B).

We also conducted a group level two-sample t-test to examine the effect of handedness on alpha peak frequency and did not find any significant difference between the left-handed and the right-handed groups ($t_{(23)}$=1.37, p=0.18).

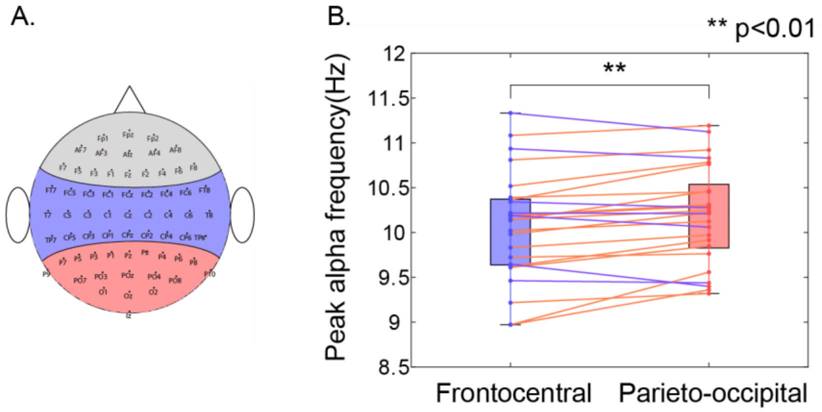

Figure 7. Peak alpha frequency of all subjects in the frontocentral and parieto-occipital region. A. Blue and red area represents the electrodes grouped into frontocentral and parieto-occipital ROI respectively. B. Blue and red colored dots represents individual alpha peak frequencies in the two ROIs respectively. Colored lines connecting the dots indicate the comparison between the two ROI. Blue lines indicate larger alpha peak frequency in frontocentral region, while red lines indicate otherwise.

3.4 Topographic specificity of alpha lateralization to auditory spatial attention focus

To quantify the amount of lateralization, GFP normalized alpha was averaged within the left and right parieto-occipital region (LROI and RROI) over the whole trial period. Alpha lateralization index was calculated by subtracting the right hemisphere ROI from the left hemisphere ROI (alpha_LROI – alpha _RROI) for each attention condition. ANOVA analysis revealed significant effect of attention focus on lateralization ($F_{(4,129)}=7.33$, $p<0.001$) but not handedness ($F_{(2,129)}=2.68$, $p=0.10$) or the interaction between handedness and attention focus ($F_{(4,129)}=0.08$, $p=0.99$). Therefore the following analysis on alpha oscillation did not separate left- and right-handed subjects.

Figure 8A shows the results of group level spatial clustering analysis. Normalized alpha power was averaged across the whole trial period. The difference between attending left (-60°) and right (60°) was used for spatial clustering analysis and plotted on a 2D topography map. A positive cluster was found on the left parieto-occipital region (p=0.007) and a negative cluster was found on the right hemisphere (p=0.009). This result is in consistency with previous literatures, where alpha oscillation power was found to decrease on the contralateral hemisphere of attended location and increases on the hemisphere ipsilateral of attended location.

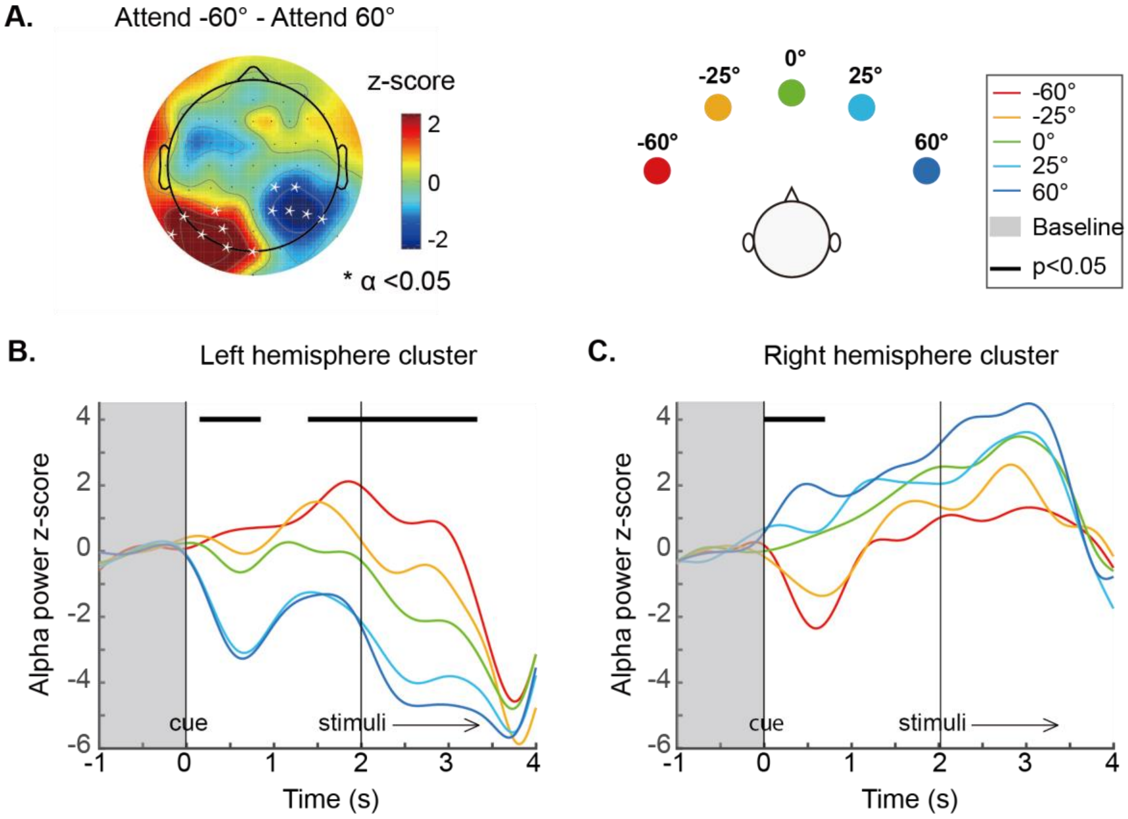

Figure 8. A. Topography of alpha power z-score difference between attend left (-60°) and attend right (60°) averaged over the whole trial (0-3.4s). Asterisks illustrate the results of the clustering analysis. Channels that showed statistical significance in the

clustering analysis were marked with asterisks. B & C. Time-courses of alpha power dynamics during the whole trial period for left hemisphere ROI and right hemisphere ROI. Alpha power was averaged within the left positive and right negative clusters resulted from the clustering analysis and plotted as time-courses for each spatial attention focus. The whole time-course was divided into 200ms time bins and averaged within each time bin. Asterisks showed the Bonferroni corrected statistical significance of alpha variation across attention focuses. Vertical black lines illustrate the onsets of auditory cue and stimuli. Shadowed areas represent the baseline period.

The electrodes within the significant clusters were selected to be ROI for temporal analysis. The positive and negative cluster revealed by clustering analysis were selected as left and right ROI respectively. Average was taken within each cluster to render the time courses of alpha power in each attention focus condition (Figure 8B & 8C). The alpha time courses reveal that alpha power in the left hemisphere increases gradually across conditions as the spatial attention focus shifts from left to right, while that of the right hemisphere decreases gradually at the same time. To quantitatively test for the significance of this effect we performed a linear regression analysis for each time bin of 200ms. Bonferroni corrected results show significant variation of alpha power when subjects are attending to different spatial locations. In the left hemisphere, significant alpha modulation occurs immediately after the cue (0.2–0.8s) and again before sound stimuli then last throughout the period sound stimuli is presented (1.6–3.2s) (Figure 8B). However, in the right hemisphere, although the trend of variation is opposite of the left hemisphere by observation and significant variation was found immediately after the cue was present (0–0.8s), no significant modulation of alpha was observed for the rest of the trial (Figure 8C).

To look at the preparatory alpha modulation, alpha power was averaged within the attentional preparation period (0s – 2s, after the cue onset before stimulus onset) for each attention focus condition and plotted onto 2D topography plots (Figure 9A). Alpha power was averaged for central electrodes and parietal-occipital electrodes (Figure 9B & C). The parietal-occipital electrodes were those revealed by the clustering analysis. Since no significant cluster was found in the central region, the central electrodes were chosen based on their location on scalp (see method 2.4.3).When attention focus shifts from left to right parietal-occipital alpha can be observed to decrease on the left hemisphere and increase on the right hemisphere (Figure 9B). On contrary, although alpha oscillation can be observed in the central cortical region on the topography plots, the lateralization pattern did not vary with attention focus (Figure 9C).

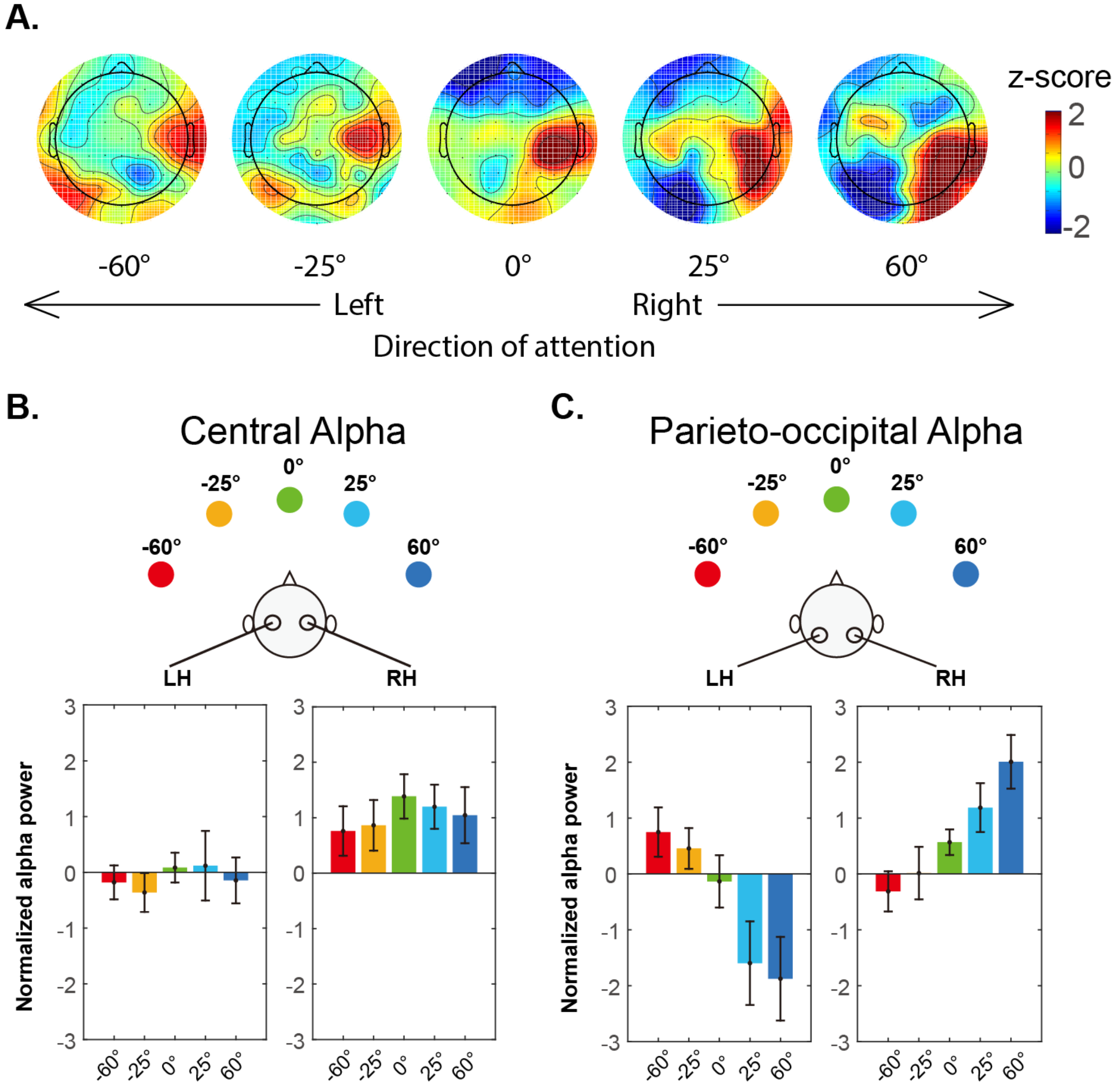

Figure 9. Alpha topographies averaged during the attentional preparatory period (-2 to 0s). A. Topographies of alpha power z-scores at five attention focuses. B. Central alpha power averaged within the left central electrodes group and the right central electrodes group. C. Parieto-occipital alpha averaged within the left and right parietal clusters revealed by clustering analysis.

4 Discussion

4.1 Behavioral performance indicates building up of top-down attention

Our results suggest that when the distractor sound appears first, endogenous (top-down) preparatory attention engagement is disrupted by the appearance of the distractor and results in lower accuracy in reporting the first syllable of the target (Figure 5). This sudden appearance of a stimuli capturing exogenous (bottom-up) attention could also explain the higher accuracy in reporting the target's first syllable when it appears before distractor. However, with time top-down attention built up and resulted in an upward trend in accuracy when listeners successfully disengage attention from the leading distractor. The building up of auditory streaming over time has been reported by previous studies (Cusack et al., 2004; Shinn-Cunningham, 2008b). The high accuracy in reporting the last digit when target is lagging, corresponding to the low accuracy in the last digit when target is leading, could alternatively be explained by that the last syllable that appears in this case (3rd syllable of distractor) is suffering from less energetic masking than the second digits and is therefore more salient to the listener.

4.2 ERP amplitudes is affected by ITD separation

Our results demonstrate that the amplitude of ERP evoked by a sound stimuli is larger when the sound is more lateralized in azimuth (larger ITD separation in this case), which is consistent with a previous study using only ITD as the spatialization cue (Dai et al., 2018) and one using ILD as the spatialization cue (Näätänen et al., 1988). However, controversial results remain about whether lateralization of sounds changes ERP amplitudes (Altman et al., 2005). Our results further suggest that the pattern of ERP amplitudes was not affected by top-down attention, with ERPs evoked by the cue (without top-down attention engagement) shows the same pattern as those evoked by the target (with top-down attention engagement). Future investigation on the ERP topography and source localization of ERP response is necessary to understand details about how ERP reflects physiological computation mechanisms of spatial cues.

4.3 Multiple alpha generators revealed by intra-individual variation in peak frequency and topographical lateralization.

Consistent with previous reports (Klimesch, 1999), we found that alpha peak frequency varied across individuals. The parietal IAF we observed has a range of 9-11.3Hz with a standard deviation of 0.6Hz. This standard deviation is lower but comparable to the previously reported 1Hz (Klimesch, 1999), which could be due to the limitation of subject number.

Interestingly, apart from inter-individual differences, we also observed intra-subject differences at different cortical regions. Alpha peak frequency in the central region is lower than that of the parieto-occipital region. Haegens et al. have examined the difference between parietal and occipital alpha peak frequencies and did not report differences (Haegens et al., 2014). The central alpha range oscillation we observed is likely to be mu rhythm whose frequency range (7.5–12.5Hz) overlaps with the alpha range and has been referred to as the somatosensory alpha (Kuhlman, 1978). There has been controversy in the frequency of mu wave compared to the alpha wave, while some claim that mu wave has a higher frequency than alpha (Van Leeuwen et al., 1978), some reports there are both a higher and a lower component (Pfurtscheller et al., 2000) and some suggests mu has a lower frequency range than alpha (Cochin et al., 1998). Our results show that 19 out of 25 subjects has a lower frequency for the central alpha than the posterior alpha. The clustering result (Figure 8A) and lateralization results (Figure 9B & 9C) show that the posterior alpha lateralizes to spatial attention focuses but not the central alpha, suggesting that posterior alpha but not central alpha plays a functional role that specific to attention focuses. Central alpha was observed to lateralize ubiquitously to the right hemisphere regardless of attention focuses. Combined, our results indicate that the alpha oscillations we observed in the central and the posterior regions are generated by different sources with potentially different functions. To understand these functions, future studies is needed by separating mu and alpha sources with source localization techniques and neuroimaging techniques with better spatial resolution. Our results further indicate that mu and alpha oscillation are likely to coexist during an active attention task and are challenging to separate due to the largely overlapping frequency range. However, most studies examining alpha define alpha as the average across a fixed band and this practice could potentially include mu rhythm and complicate the interpretations.

As expected we did not find any difference in alpha peak frequency across different attentional conditions. While previous studies have shown that different task loads could affect the IAF (Basar, 2012; Haegens et al., 2014), the different attention conditions in our task are at the same difficulty level as suggested by the behavioral performance results.

4.4 Topographic specificity of alpha power lateralization to auditory spatial attention focus

We showed that when auditory attention is covertly oriented to a spatial location alpha oscillation forms lateralization patterns topographically specific to the attention focuses. In line with the previous studies (Frey et al., 2014; P. Sauseng et al., 2005; Strauß et al., 2014; Thut et al., 2006; Wöstmann et al., 2016a), we observed that alpha in the parieto-occipital region desynchronizes on the contralateral hemisphere to the attended hemifield and synchronizes on the ipsilateral hemisphere (contralateral hemisphere of the unattended hemifield). Our study extended the attention focuses beyond simply left and right to examine five different attention focuses along the azimuth. Our results revealed that lateralization of posterior alpha depends on the laterality of the attention focus. The more lateralized the attention focus is, the more lateralized the alpha pattern is. These results agree with the previous results (Rihs et al., 2007a) that observed retinotopical specificity of alpha during visuospatial attention, where alpha distribution pattern was found to be less correlated to the central attention focus condition when the azimuth of attention focus increases. Also in line with (Rihs et al., 2007a), we found the alpha lateralization pattern along azimuth to be symmetric to midline. The direction of alpha lateralization could be explained by the inhibition theory (Klimesch et al., 2007b) that regards alpha as a suppression mechanism that gate the processing of sensory inputs from contralateral hemifield. However, the magnitude of lateralization implies more complicated underlying modulation mechanism that requires further exploration. Based on the assumption that alpha functions as an active suppression mechanism, a possible explanation is that when attention is focused in the center, probabilities of distraction coming from left and right side are equal, and thus suppression on both hemispheres are counterbalanced. When attention focus is at the extreme laterals, all distraction locations would be at one side of the focus, therefore the suppression on the ipsilateral hemisphere would overpower the contralateral counterpart.

Moreover, our results support the supramodal theory by showing that audiospatial attention demonstrates similar topographical specificity as visuospatial attention, suggesting that auditory and visual spatial attention share the same parietal network. There has been numerous studies on the retinotopical specificity of the visual spatial attention network (Golomb et al., 2008; Silver et al., 2005) but very few on auditory spatial attention network. In future experiments, it would be interesting to include more attention focuses not limited to the frontal plane with zero elevation. Also, source localization of the alpha activities would be useful to review the cortical generator of alpha and whether it has the same topographical organization as the visual spatial attention network.

4.5 Hemispheric asymmetry in alpha modulation by spatial attention

We observed hemispheric asymmetry in posterior alpha modulatation during auditory spatial attention. The time-courses of the alpha modulation revealed that for left hemisphere, alpha modulation lasted 600ms after auditory cue onset (Figure 8B & 8C). Similar results were observed for the right hemisphere. However, in the left hemisphere but not the right hemisphere significant modulation occurs again 400ms prior to the stimulation onset and then lasted throughout the stimulation period. The brief post-cue period of lateralization could contain part of the phase-locked response to the auditory cues, which were also spatially lateralized; while the post-cue pre-stimulus period of modulation contains mainly induced response. We also noticed that the amount of modulation is larger in the left hemisphere than the right hemisphere. Compared to the left hemisphere, right hemisphere alpha showed less desynchronization when attending to the contralateral spatial hemifield, therefore led to a weaker lateralization pattern. The asymmetries we observed are similar to the results of a previous visuospatial attention study (Ikkai et al., 2016b). A previous tactile spatial attention study also reported larger alpha modulation in the left hemisphere than the right hemisphere (S. Haegens et al., 2011). The hemispheric difference in alpha modulation supports our hypothesis based on the right hemisphere dominance theory, and our results are in line with previous results by (Huang et al., 2014) that the left hemisphere has a more lateralized connectivity pattern with the contralateral AC. The right hemisphere on the other hand has a bilateral representation of both ACs and therefore shows a less lateralized pattern.

Although we included left-handed subjects in the study with the intention to study effects of atypical hemispheric asymmetry to alpha modulation, we did not find any significant difference in alpha modulation between left- and right-handed listeners. However, this result may have been confounded by the fact that a low percentage (1 in 10) of left-handers have atypical hemispheric dominance (Knecht, 2000). There may not be enough subjects with atypical hemispheric dominance in our left-handed group to show potential effects. Future studies with a prescreen procedure of the hemispheric dominance could possibly illicit more informative results (Cai et al., 2013).

Conclusion

We studied the underlying neural mechanisms when auditory attention is focused at five azimuth locations (-60°, -25°, 0°, 25°, 60°). Our results of spectral analysis showed different alpha peak frequencies at different cortical regions, suggesting a central mu rhythm source and parieto-occipital alpha source. Unlike central alpha, parieto-occipital alpha lateralizes in a topographically specific manner to the five spatial attention focuses along the azimuth. Furthermore, we observed asymmetry between left and right hemisphere in alpha power modulation. While left hemisphere shows graduate alpha power modulation across attention focuses, the right hemisphere showed a less obvious and statistically insignificant modulation trend. Combined our results fit predictions from the supramodal attention network hypothesis and Neglect model.

CHAPTER TWO

Differences in cortical processes during spatial and non-spatial divided auditory attention

Yuqi Deng[1], Barbara Shinn-Cunningham[2], Inyong Choi[3]

Statement of permission:

This manuscript has been prepared for submission to a peer-reviewed journal.

[1] Department of Biomedical Engineering, Boston University, 610 Commonwealth Ave., Boston, MA, USA, 02215;

[2] Carnegie Mellon Neuroscience Institute, Carnegie Mellon University, Pittsburgh, PA 15289;

[3] Department of Communication Sciences and Disorders, University of Iowa, Iowa City, IA, 52242;

Abstract

Social settings require us to not only focus attention on one talker, but also to monitor the scene and reorient rapidly when a new talker triggers interest. However, few studies have explored the neural underpinning of dividing attention across a scene, especially while such strategies operate on different acoustic features. Here with electroencephalography we investigated differential cortical dynamics during focused and divided attention when listening to competing speech streams separated by spatial or non-spatial features. In each trial, listeners were visually cued to either focus attention on an initial Target or divide attention to monitor a later “Supertarget” that appears with a 50% chance. In spatial blocks, the Supertarget was defined by its direction and monitored by attending to this additional direction. In non-spatial blocks, the Supertarget was defined by its talker identity and monitored by attending to this additional voice. Behaviorally, the ability to report the initial Target was impaired when attention was divided to monitor two directions but not two voices. In both forms of attention, focusing on the initial Target led to lateralized alpha oscillation power. However, alpha lateralization disappeared during dividing spatial attention but not non-spatial attention, corresponding to monitoring the additional direction but not the additional voice. These results provided insights of differences in behavioral strategies and neural mechanisms when auditory attention is divided based on spatial and non-spatial acoustic features. Monitoring multiple auditory objects based on non-spatial features is more intuitive and requires less top-down attentional efforts than on spatial features.

Key words: auditory; divided attention; spatial, non-spatial attention; EEG; ERP; alpha oscillation;

Significance statement

Auditory selective attention in social settings is often divided to monitor multiple sound streams that are both of interest, which may increase cognitive load. This study presents a novel divided attention task to allow direct comparison of cognitive load and cortical neural dynamics between spatial and non-spatial attention. As our results suggest that it is easier and more natural to detect the appearance of a particular talker from any location than an unknown talker from a particular direction, we accelerate understanding of cortical mechanisms underlying these two forms of attention.

1 Introduction

In social settings where conversations ebb and flow, in addition to focusing attention on a target speaker, it is also important to monitor other speakers to actively manage conversations by directing and redirecting attention. Such divided attention may demand extra cognitive load (Corbetta et al., 1991; Triesman, 1969), which forms a neural underpinning of increased listening effort and fatigue during conversations in social settings (Madden and Plude, 1993; Strauss and Francis, 2017). However, it remains unclear what neural mechanisms underlie top-down cognitive strategies for divided auditory attention, especially when such strategies operate on different acoustic features, including spatial acoustic features and non-spatial features (e.g., Darwin et al., 2003; Kidd et al., 2005; Maddox and Shinn-Cunningham, 2012; Bressler et al., 2014).

Previous studies that employed stationary attention tasks where target and masker streams are fixed throughout trials unveiled neural substrates including increased peripheral sensitivity to attended signals (Posner et al., 1980), enhanced cortical responses to attended stimuli (Choi et al., 2013; Hillyard et al., 2007; Woldorff et al., n.d.); (Alho et al., 1987; Degerman et al., 2008), and hemispheric lateralization of cortical oscillations (especially during *spatial* selective attention). Specifically, alpha band (8-14Hz) oscillation has been demonstrated to be modulated by top-down spatial attention. Alpha synchronization on the hemisphere ipsilateral to the attended side and alpha desynchronization contralateral to the attended side have been observed in both vision (e.g. Worden et al., 2000a; Kelly, 2006; Thut, G., Nietzel A, Brandt SA, 2006) and audition (Frey et al., 2014; Klatt et al., 2018; Mehraei et al., 2018; Wöstmann et al., 2016a). Such alpha oscillation is thought to be supramodal (Snigdha Banerjee et al., 2011; Kerlin et al., 2010) and plays the functional role of inhibiting sensory encoding of inputs from a contralateral to-be-ignored location (Jensen and Mazaheri, 2010a; Klimesch, 2012a). However, it is controversial whether parietal cortex process spatial and non-spatial attention differentially (Arnott et al., 2004; Bidet-Caulet and Bertrand, 2005; De Santis et al., 2007; Degerman et al., 2006; Lee et al., 2013) and therefore whether this fronto-parietal inhibition mechanism is also employed by non-spatial attention. While some vision features-based studies observed alpha modulation (Snyder and Foxe, 2010), some did not (H A Slagter et al., 2016).

The goal of this study is to find neural substrates of divided auditory attention, focusing on the differences in cognitive load between spatial and non-spatial attention. In visual domain, dividing visual attention requires more cognitive resources (Madden and Plude, 1993) and thus results in insensitivity for discriminating subtle differences in stimuli (Corbetta et al., 1991). Similarly in auditory domain, due to the increased cognitive load, divided auditory attention is more prone to performance decline caused by aging than focused attention (Wild-Wall and Falkenstein, 2010; Getzmann et al., 2016). Larson and Lee (2014) showed different cortical networks involved during auditory attention switching based on spatial and non-spatial features. However, none of these studies discussed neural dynamics that require divided attention to multiple competing sound sources, where listeners have to monitor a potential new sound while attending to an existing target.

We hypothesize that differential behavioral strategies during a divided attention task reflects critical differences in neural mechanisms between attending spatially and non-spatially. Spatial cues are strong for selective attention but not syllabic level grouping (for review: Shinn-Cunningham et al., 2017). Therefore, in terms of behavioral performance, we hypothesize that when divided spatial attention spotlight covers wider areas rather than a distinct location, spatial separation within the attention spotlight would weakly benefit syllabic level grouping. In terms of neural processes, we hypothesize that divided spatial attention engages more domain-general (i.e. not auditory-specific) dorsal network including the modulation of parietal alpha but not non-spatial attention. We developed a dividing attention paradigm to directly compare behavioral strategies, cognitive load, and underlying neural dynamics involved in spatial and non-spatial auditory selective attention, expecting coherent differences found from behavioral and neural results.

2 Method

2.1 Subjects

Thirteen subjects (six females, 23–37 years of age) participated in the experiment. All subjects had normal hearing (hearing thresholds better than 20 dB at pure tone frequencies between 250 Hz and 8 kHz). All gave informed consent as approved by the Boston University Institutional Review Board. One subject performed the task at chance levels; their data were therefore excluded from further analysis.

2.2 Stimuli

Each trial consisted of three streams of sound, each of which contained a string of randomly selected syllables chosen from /ba/, /da/, and /ga/ (chosen because their plosive onsets elicit strong event-related potentials or ERPs in EEG). Three native English speakers recorded the syllables using an AudioTechnica AT4033 large diaphragm condenser microphone (Audio-Technica U.S. Inc., Stow, OH) in an Eckel sound proof booth. The three talkers were a male with relatively high pitch (F0 of 125Hz), another male with relatively low pitch (F0 of 91Hz), and a female (F0 of 189Hz). Individual syllables for each talker were recorded in isolation, then concatenated to form three- or five-syllable-long random streams.

In each trial, the first stream to begin playing was the Target stream, which consisted of three randomly selected syllables. The second stream, known as the Distractor, started 150 ms after the Target, and consisted of five randomly selected syllables. Finally, on 4 out of 5 trials (randomly selected; see Task description in Section 2.3), a three-syllable-long Supertarget appeared; if it appeared, its timing was random, equally likely to begin 1 s or 1.5 s after the Target (start times that fall either between the 2nd and 3rd Target syllables or after the final Target syllable, respectively). To discourage listeners from using event timing to direct attention, streams were not isochronous; instead, the timing between adjacent syllables within a stream alternated between 500ms and 750ms, with the initial interval chosen randomly on each trial. With these stream timings, the onsets of the individual syllables across all streams were separated sufficiently to allow the event-related potentials (ERPs) evoked by each syllable to be resolved in time.

Both Target and Distractor were presented diotically, and thus were perceived at midline. The Target stream always contained syllables spoken by the high-pitched male talker, while the Distractor stream syllables were spoken by the low-pitched male talker. The interaural time difference (ITD) of the Supertarget differed in different blocks. In Spatial Attention blocks, the Supertarget always had an ITD of 700 µs leading in the left ear (perceived as coming from left of midline), with syllables spoken either by the high-pitched male talker or the female talker (randomly selected on each trial). In Talker Attention blocks, the Supertarget in a given trial had a randomly selected ITD of either 700µs (from the left) or 0 µs (from midline); its syllables were always spoken by the female talker.

2.3 Task

Here we adopt a switch attention task with 2×2 factorial design of attentional strategies, where human listeners were cued to fix or switch auditory attention based on either spatial location or talker identity, since in everyday settings, focusing attention on a specific talker is natural and the most often deployed non-spatial attention strategy. We would refer to our non-spatial auditory attention as Talker attention.

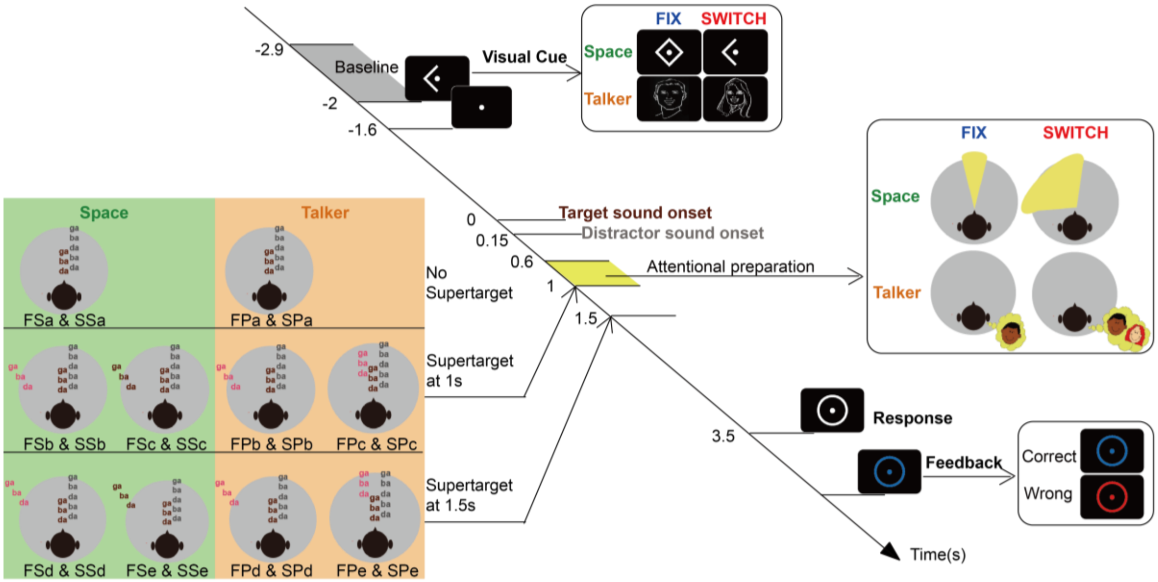

Figure 10. Paradigm design. We used a task where subjects are visually cued to focus or switch attention based on spatial (diamond/left arrow) or pitch (male/female picture). A visual cue is present at the beginning of the trial. It lasts for 400 ms and is followed by a fixation dot. 2s after the visual cue onset the target sound (three syllables in high pitched male voice) starts playing, followed by the distractor sound (five syllables in low pitched male voice) that starts 150 ms later. The Supertarget (three syllables) may not occur or may occur at 1 s or 1.5 s after target onset. In space attention tasks the possible Supertarget appears from the left (700µs ITD) and the voice alternates between female voice and high-pitched male voice randomly across trials. In pitch attention tasks the possible Supertarget is a female voice and its location alternates between center (0µs ITD) and left (700µs ITD) randomly across trials. Subjects report the sequence of syllables as indicated by the previous visual cue after a response cue that appears 3.5 s after target onset. As soon as response is reported a

feedback cue appears indicating whether the answer was correct or not. The conditions are labelled in the following format: XYz. First letter X is either FOCUS (F) or DIVIDE (S); Second letter Y is either space (S) or pitch (P); Subscript z is a, b, c, d, e, describing details of stimulus.

A visual cue at the start of each trial informed listeners to either FOCUS attention on the Target or to be prepared to DIVIDE attention to the Supertarget— if it appeared (see Figure 10). The cue appeared 2 s before the onset of the acoustic stimuli in a trial; after 400, it switched to a fixation dot that remained until the end of the trial. Listeners were asked to hold their gaze on the fixation dot throughout each trial. The auditory mixture of Target, Distractor, and Supertarget (if present) then played. Exactly 3.5 s after the start of the auditory mixture, the response period began (indicated by a white circle appearing around the fixation dot). Listeners gave their responses in the response period by typing the perceived three-syllable sequence on the keyboard. Correct answer feedback was provided after the response period by a change in color of the circle (green for correct, red for incorrect).

On FOCUS trials, listeners were instructed to report the sequence of Target syllables, ignoring all other streams. On DIVIDE trials, subjects were instructed to begin by attending to the Target and to report its content if no Supertarget appeared; however, if the Supertarget appeared, they were asked to report its syllable sequence, instead. At the end of each trial, subjects gave their responses by typing the heard syllable sequence on the keyboard.

Trials were presented in blocks that differed in whether subjects were to focus attention based on lateral position (Spatial Attention) or on talker identity (Talker Attention). In FOCUS trials, regardless of the block, listeners had to focus attention on the Target, which was always diotic, and always with syllables spoken by the high-pitched male talker (indicated by a visual cue that was a diamond in the Spatial Attention blocks or a male icon associated with the high-pitched male in the Talker Attention blocks; see Figure 10). In DIVIDE trials during Spatial Attention blocks, the Supertarget was always from the left (indicated by a leftward arrow visual cue; see Figure 10), but could be either the same high-pitched male talker as the Target talker (differentiable from Target by spatial feature only) or the female talker (differentiable from Target by both spatial and voice feature). In DIVIDE trials during Pitch Attention blocks, the Supertarget was always the female talker (indicated by a female icon visual cue; see Figure 10), but could be either be from the same midline location as the Target (differentiable from Target by voice feature only), or to the left (differentiable from Target by both spatial and voice feature). Thus, in Spatial Attention blocks, the Supertarget voice varied and listeners had to divide attention in DIVIDE trials between the midline Target and a potential Supertarget from the left; in Talker Attention blocks, the Supertarget location varied and listeners had to divide attention in DIVIDE trials between the high-pitched male Target and a potential Supertarget spoken by the female target.

With this design, there are physically identical FOCUS and physically identical DIVIDE trials in both Spatial Attention and Talker Attention blocks. For instance, trials where no Supertarget appeared were identical in both blocks, and listeners had to report the Target in these cases. In DIVIDE trials, both Spatial Attention and Talker Attention blocks included trials with a leftward Supertarget that has syllables spoken by the female talker. These choices allowed us to directly compare behavioral performance and cortical responses when subjects performed the same task and were asked to report the same stream, but may have employed different attentional strategies due to the differences in expectation.

Each subject performed 22 blocks of 40 trials each: 11 blocks of Spatial Attention, and 11 blocks of Talker Attention. Each block contained 20 FOCUS and 20 DIVIDE trials, in random order. Of the 20 trials for each instruction type, 4 had no Supertarget and the other 16 were made up of 4 repetitions of each combination of Supertarget parameters. Specifically, in Spatial Attention blocks, Supertargets were presented for the four combinations of potential talker (female or high-pitched male) and delay (1 s or 1.5 s after the Target), while in Talker Attention blocks, Supertargets were presented for the four combinations of potential locations (leftward or midline) and delay. With this design, each block included four trials of exactly the same instruction (FOCUS or DIVIDE) and physical stimulus, leading to a total of 44 repetitions of each kind of trial across the 11 blocks. Although equal numbers of trials were presented for each possible syllable rhythm, trials with different rhythms were averaged together to achieve reasonable reliability. Only trials on which subjects answered correctly were used in subsequent EEG analysis of ERPs and alpha oscillations. Across subjects and conditions, the number of correct trials ranged from 21 to 44.

2.4 EEG acquisition and analysis

EEG data was recorded with 64-channel Biosemi ActiveTwo system in an Eckel sound-treated booth while subjects performed the experiment. Two additional reference electrodes were placed on the mastoids. Stimulus presentation was controlled by Matlab (MathWorks, Natick, MA) using the Psychtoolbox 3 extension (Brainard, 1997). Both the EEGlab toolbox (Delorme and Makeig, 2004a) and the Fieldtrip toolbox (Oostenveld et al., 2011) were used for EEG analysis. Raw EEG data were referenced against the average of the mastoid channels and down-sampled to 256 Hz.

ERP analysis was performed from recordings from Channel Fz. A 2048-point FIR filter was used to bandpass filter the raw recordings from 2-20 Hz. ERPs were converted to z-scores; we baseline corrected by subtracting the mean and then divided by the standard deviation of the baseline period, defined as the 500-ms period prior to the visual cue. To reduce artifacts, the trials whose instantaneous amplitude was in the top 10% of all trials in each condition were thrown out. Within each condition, the EEG recordings from the trials on which listeners answered correctly were discarded; ERPS were derived from the average of only the correct-answer trials.

We were interested in examining how ERPs evoked by the Target and Supertarget were modulated by dividing attention (i.e., comparing ERPs in FOCUS versus DIVIDE trials). To analyze effects of attention on the first Target-evoked ERP, we combined all trials with the same Type of Attention (Spatial or Talker) and Attentional Instruction (FOCUS or DIVIDE), ignoring what came later (i.e., collapsing across trials with no Supertarget and with different kinds of Supertargets). We then found the amplitude of the N1 response evoked by the first syllable of the Target as the maximum negative peak amplitude within the time window 50-150 ms after its onset. To analyze effects of attention on the first Supertarget-evoked ERP, we classified trials based on the combination of Type of Attention (Spatial or Talker), attentional instruction (FOCUS or DIVIDE), and Supertarget onset time (1 s or 1.5 s after the Target onset). To ensure we had enough repetitions for the Supertarget ERP analysis, we collapsed across different Supertarget talkers in the Spatial Attention blocks, and across different Supertarget directions in the Talker Attention blocks. We then found the amplitude of the N1 response evoked by the first syllable of the Supertarget as the maximum negative peak amplitude within the time window 50–150 ms after its onset.

To analyze alpha oscillation activity, we used a 2048-point FIR filter to bandpass filter raw EEG responses between 8-14 Hz. We took the Hilbert transform of the bandpass-filtered data; the absolute value of the Hilbert transfer was used to estimate the instantaneous alpha oscillatory amplitude. The alpha power time course was baseline-corrected using the mean value over the 900ms preceding the visual cue onset (the baseline period before the appearance of the visual cue was used to avoid any possible differences in alpha induced by different visual cues). To reduce artifacts, the trials whose instantaneous alpha power was in the top 10% across all trials were thrown out. Alpha power at each electrode was normalized by subtracting the mean alpha power of all electrodes and dividing by the global field power (GFP: Skrandies, 1990; Murray et al., 2008) which yielded spatial z-scores for each electrode. Based on the topography of alpha responses from previous studies, we selected electrodes P5 and P6 (over left and right parietal cortex, respectively) a priori as the channels of interest (Snigdha Banerjee et al., 2011; Kelly, 2006b; Worden et al., 2000a). We extracted the normalized alpha power in these two channels as a function of time from the start of the visual cue up to the time at which the Supertarget could first appear, then bandpass filtered from 0.1–10 Hz and averaged across subjects (visualized in Figure 14).

Based on results of the clustering analysis, we chose the time periods of interest as those that contained a significant or trending significant time cluster in *both* the left and right parietal channels in either the Spatial Attention or Talker Attention conditions. To avoid filtering artifacts associated with the response evoked by the Target and the Supertarget, we excluded 50 ms periods just before the Target and Supertarget onset (i.e., -50ms to 0ms and 950 ms to 1000 ms) in the final time clusters analyzed. Normalized alpha power was averaged across these clusters, and then averaged across all trials for each of the four attentional strategies. Finally, we computed the distribution of alpha power averaged over these time clusters for each sensor to create a 2D topography map.

2.5 Experimental design and statistical analysis

We conducted dependent samples parametric and non-parametric statistical tests for direct comparison between conditions. Due to our modest sample size (N=12), the normality of samples were checked by Anderson-Darling normality test before performing a parametric statistical analysis (Razali and Wah, 2011). A dependent-samples t-test was used for statistical comparison when the normality test did not indicate deviation from the normal distribution (Anderson-Darling test: $p>0.05$). If the normality assumption did not hold (Anderson-Darling test: $p<0.05$), a nonparametric Wilcoxon rank-sum test was performed instead for pairwise comparisons (Wilcoxon, 1945). All pairwise tests were two-tailed with a significant threshold of 0.05.

To compare the time course of differences between FOCUS and DIVIDE attention, we performed a group-level nonparametric statistical test implemented in Fieldtrip. The analysis used dependent samples t-tests to compare FOCUS and DIVIDE trials at each time sample, then clustered together adjacent time samples with similar values. The statistical significance of the cluster is then computed by performing a permutation test with 1000 Monte Carlo iterations (with replacement) using an alpha level set to 0.05 (Lim et al., 2015; Maris and Oostenveld, 2007).

3 Results

3.1 Behavioral results

Performance in both Spatial Attention and Talker Attention conditions was generally good, well above chance level of 33%, for both FOCUS and DIVIDE conditions and both with and without a Supertarget present (see Figure 12&13). When no Supertarget was present, performance was similar in Spatial Attention and Talker Attention conditions for the FOCUS trials (Figures 11). In all conditions where subjects reported the Target stream regardless of the timing of the potential Supertarget, performance is best for the first syllable of the Target and decreases from syllable to syllable (Figure 11). Generally in FOCUS trials listeners were the most accurate in reporting the Target when the to-be-ignored Supertarget did not appear, and they performed the worst when Supertarget disrupted the Target at 1 s. This pattern of performance was observed in all conditions regardless of the number of differential features the Supartarget had (Figure 11). In fact, the number of differential features the Supertarget had did not have a significant influence on behavioral performance in either Spatial or Talker Attention (Figure 11).

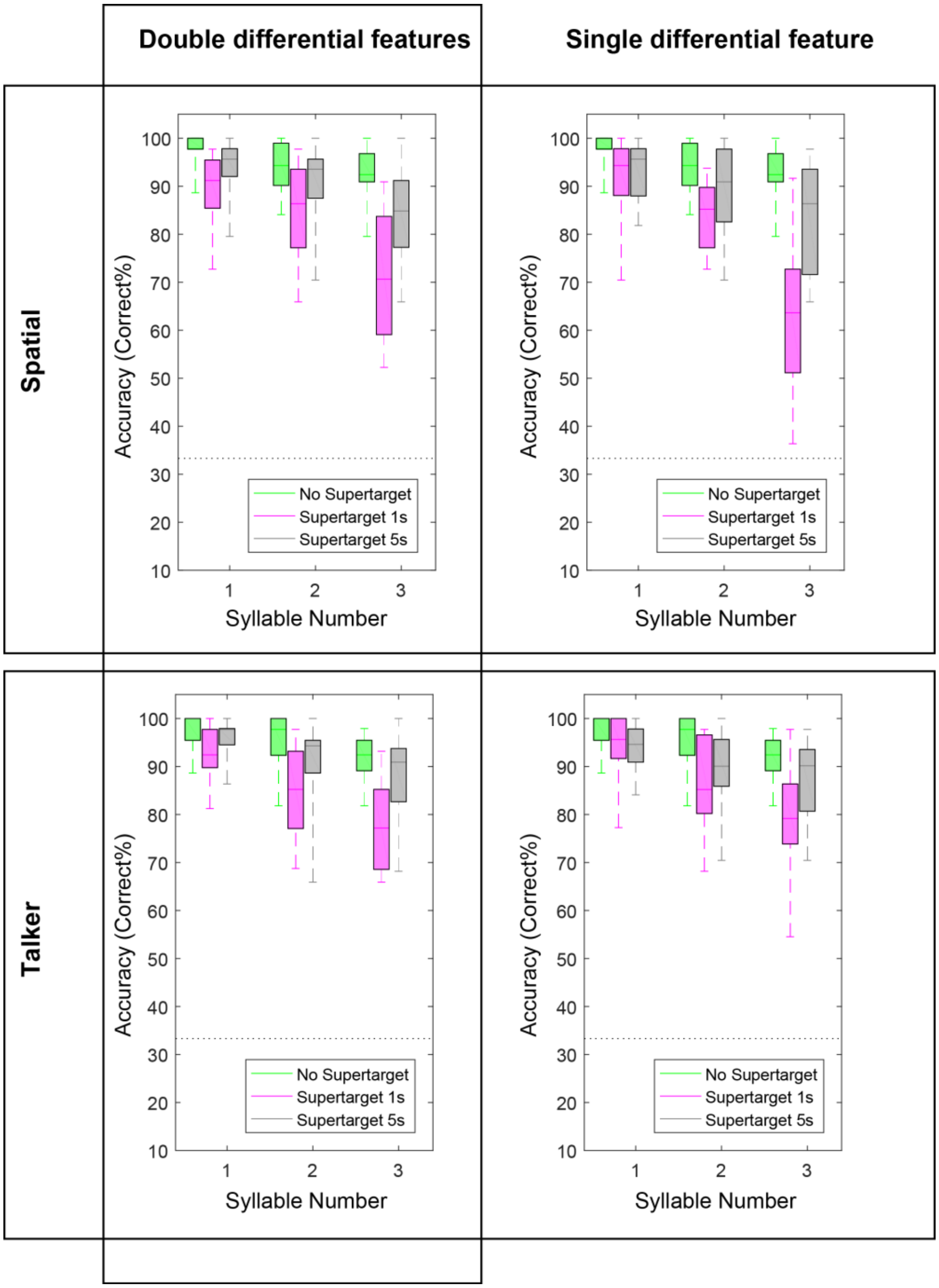

Figure 11. Behavioral performance of FOCUS trials when Supertarget did not appear or appear at different timings. Comparisons could be made within Spatial or Talker Attention conditions, or between Spatial and Talker Attention in case where Supertarget has double differential features from the Target.

3.1.1 Effect of dividing attention

To examine the effects of dividing attention, we compared behavioral performance for FOCUS and DIVIDE trials in which the Supertarget did not appear (Figure 12). For both types of trials without a Supertarget, the listener simply had to report the content of the Target. Thus, we compared trials with identical physical stimuli for which the correct response was the same, but in which listeners used different attentional strategies: they either focused on the Target or listened to the Target but were prepared to switch attention to a Supertarget.

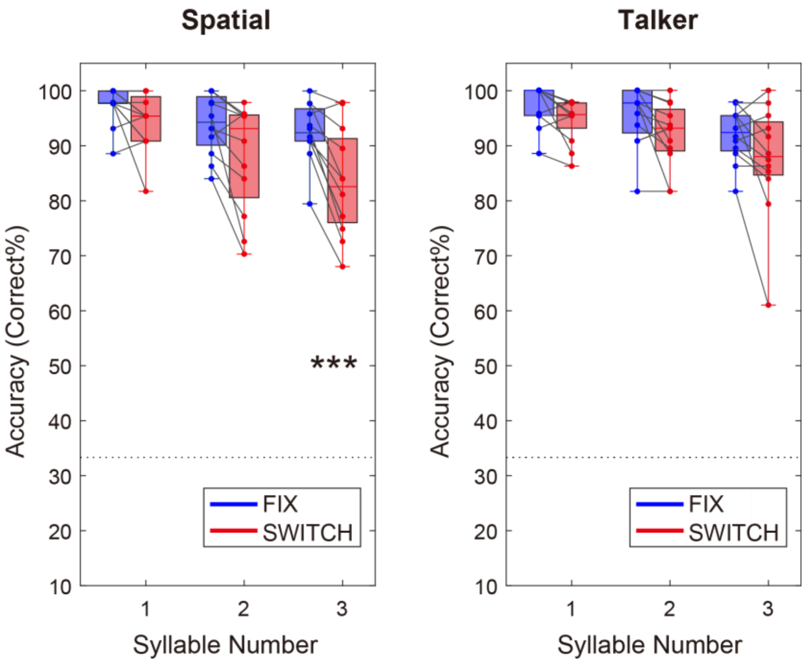

*Figure 12. Behavioral performance of FOCUS and DIVIDE trials when Supertarget did not appear. Asterisk indicates statistical significance (*p<0.05; **p<0.01; ***p<0.001).*

The percentage of correct responses were generally lower in DIVIDE than in FOCUS trials. In the Space Attention trials, a group-level pairwise t-test or Wilcoxon rank-sum test revealed that this difference reached significance for the third syllable ($t_{(11)}$=5.27, p=0.0003) but not for the first syllable ($z_{(11)}$=1.51, p=0.13) and the second syllable ($z_{(11)}$=1.42, p=0.16). In Talker Attention trials, the listening strategy had no significant effect on performance for any syllables (1st syllable: $z_{(11)}$=1.85, p=0.064; 2nd syllable: $z_{(11)}$=1.66, p=0.097; 3rd syllable: $z_{(11)}$=2.07, p=0.063). Our results suggest that for Spatial Attention there is a performance cost associated with monitoring the sound mixture for the appearance of a Supertarget, even if it does not appear. However, in Talker attention this effect of dividing attention is not statistically significant.

3.1.2 Effect of a to-be-ignored Supertarget

To examine the effects of to-be-ignored distractions, we compared behavioral performance in trials where the Supertarget appeared 1 s after the Target (before the final Target syllable) and when it appeared 1.5 s after the Target (after the final Target syllable) for all of the FOCUS trials (See Figure 13). In all the FOCUS trials listeners know that they should maintain attention on the Target. The potential Supertarget could have one differential feature from the Target or both differential features (See task description in section 2.3). We compared performance between Spatial and Talker Attention when the Supertarget differs from the Target in both features, because in these trials the Spatial and Talker Attention conditions had the same physical stimuli and thus directly comparable.

There was no significant difference in performance between Spatial and Talker Attention in the first two syllables regardless of the timing of the Supertarget (when Supertarget appeared at 1 s: 1st syllable, $z_{(11)}$=1.01, p=0.31; 2nd syllable, $z_{(11)}$=0, p=1;

when Supertarget appeared at 1.5 s: 1st syllable, $t_{(11)}$=1.87, p=0.88; 2nd syllable, $t_{(11)}$=0.20, p=0.85;). However, for the third syllable, which is the syllable interrupted by the Supertarget at 1 s, subjects performed significantly better in Talker Attention than in Spatial Attention ($t_{(11)}$=2.38, p=0.037). No difference of the third syllable were observed when it appeared at 1.5 s after the Target and did not interrupt the Target ($t_{(11)}$=0.37, p=0.71).

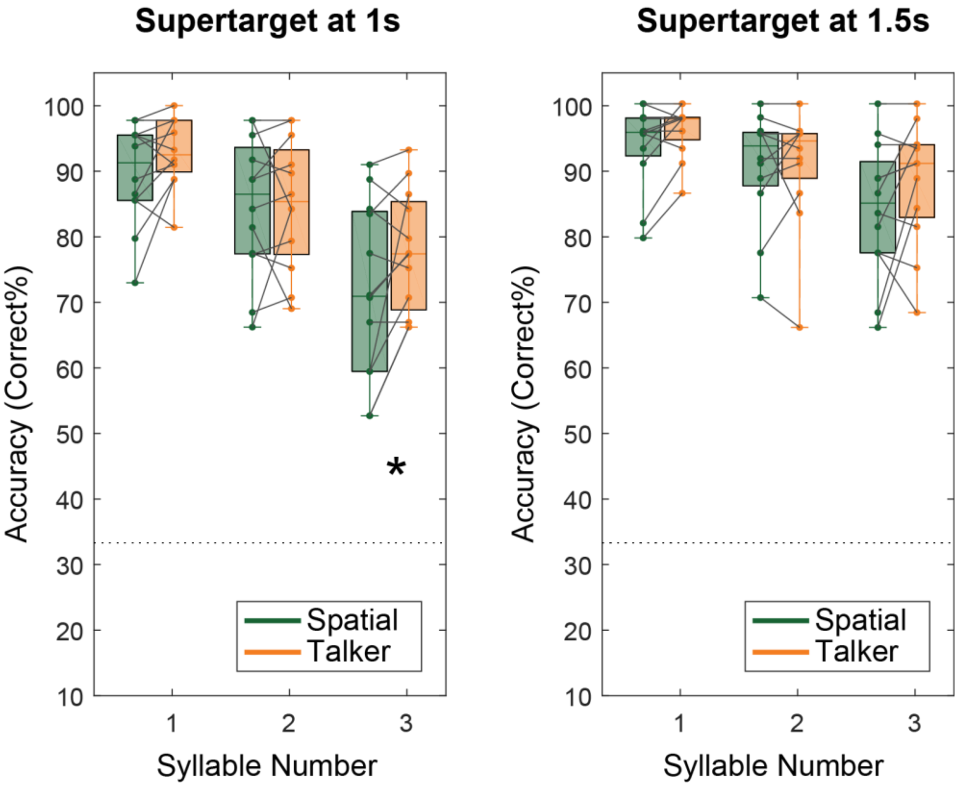

*Figure 13. Behavioral performance of Spatial and Talker attention trials when Supertarget appeared at 1 s and 1.5 s after Target. Asterisk indicates statistical significance. Asterisk indicates statistical significance (*p<0.05; **p<0.01; ***p<0.001).*

3.2 EEG results

3.2.1 Effects of attentional state on evoked neural responses

We analyzed the magnitude of the N1 response to the onset of the Target stream, which is the first sound in the acoustic mixture. We found no significant differences in the Target N1 magnitude for FOCUS versus DIVIDE trials for either Spatial Attention ($t_{(11)}$=0.74, p= 0.47) or Talker Attention ($t_{(11)}$=0.68, p= 0.51).

We then analyzed the N1 response evoked by the onset of the Supertarget separately when it was 1 s after the Target (before the final Target syllable) and when it was 1.5 s after the Target (after the final Target syllable). In both Spatial Attention and Talker Attention conditions, the N1 for the Supertarget at the 1 s delay was larger in the DIVIDE than in the FOCUS trials (Spatial Attention: $t_{(11)}$=2.78, p=0.018; Pitch Attention: $t_{(11)}$=3.56, p=0.004), consistent with the Supertarget evoking a greater response when listeners are trying to attend to it than when they are trying to ignore it (see Figure 14a & 14e). We calculated the strength of attentional modulation of the Supertarget as the amplitude of the N1 evoked in the DIVIDE condition minus the amplitude of the N1 evoked in the FOCUS trials. The strength of attentional modulation of the Supertarget was not significantly different for Spatial Attention and Talker Attention ($t_{(11)}$=0.43, p =0.68). In contrast, the N1 response evoked by the onset of the Supertarget that started 1.5 s after the Target was not significantly different for FOCUS and DIVIDE trials during either Spatial Attention ($t_{(11)}$=0.79, p=0.45) or Talker Attention ($t_{(11)}$=0.69, p=0.50; i.e., there was no attentional modulation of the Supertarget that started after the Target ended. See Figure 14b & 14f).

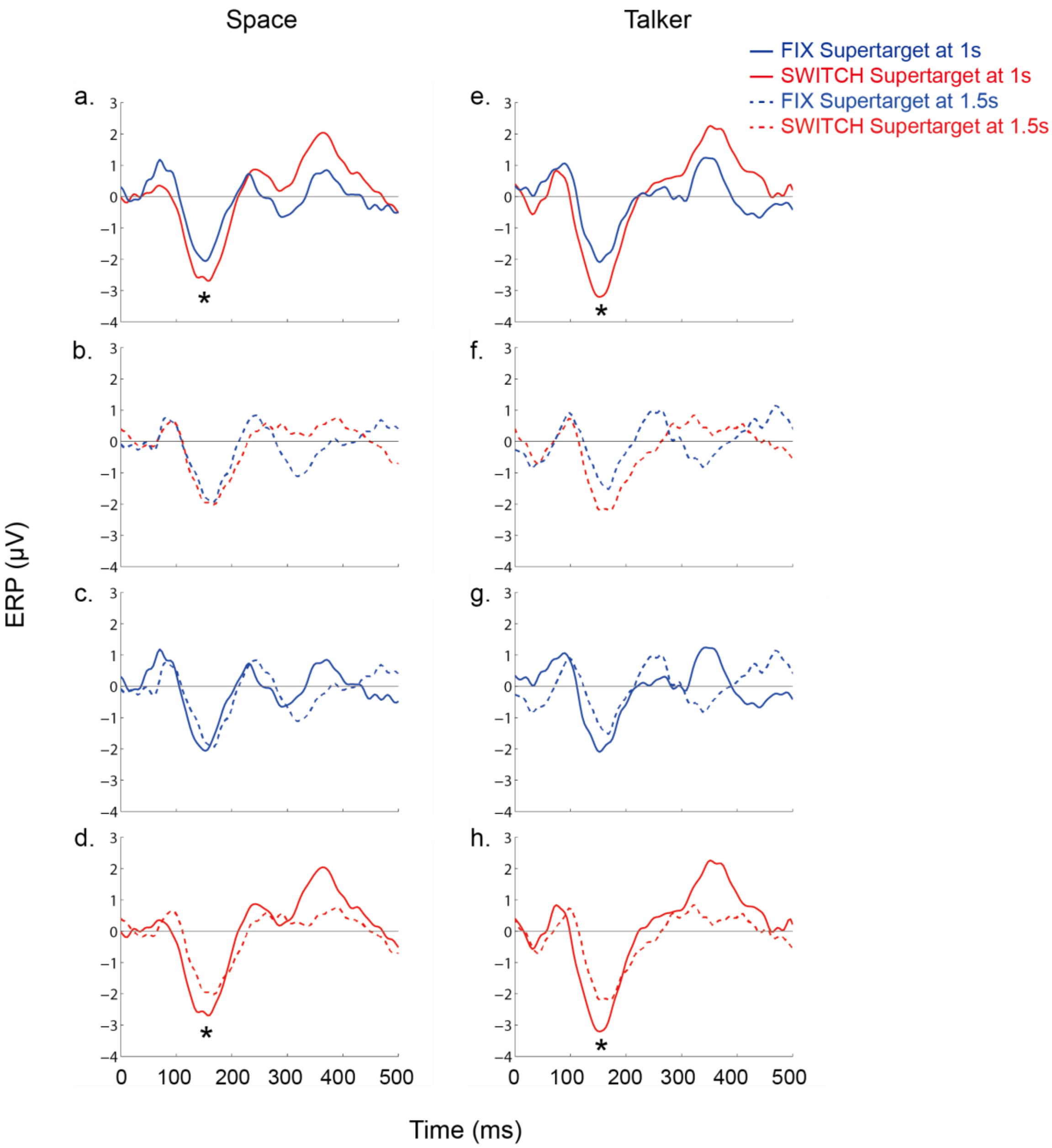

Figure 14. ERP evoked by the first syllable of the Supertarget at channel Fz. The left column shows results for Space Attention and the right column for Talker Attention. A-B) Supertarget at 1 s compared between FOCUS trials (FSbc: average of FSb and FSc) and DIVIDE trials (SSbc: average of SSb and SSc). b & f) N1 amplitude evoked by Supertarget onset at 1.5 s compared between FOCUS trials (FSde: average of FSd and

*FSe) and DIVIDE trials (SSde: average of SSd and SSe). c & g) N1 amplitude comparison between Supertarget onset at 1 s in FOCUS trials (FSbc: average of FSb and FSc) and Supertarget onset at 1.5 s in FOCUS trials (FSde: average of FSd and FSe). d & h) N1 amplitude comparison between Supertarget onset at 1 s in DIVIDE trials (SSbc: average of SSb and SSc) and Supertarget onset at 1.5 s in DIVIDE trials (SSde: average of SSd and SSe). Asterisk indicates statistical significance in N1 amplitude comparison (*p<0.05; **p<0.01; ***p<0.001).*

For both the FOCUS and the DIVIDE trials, we then compared the N1 evoked by the Supertarget when it appeared at 1 s (before the final Target syllable) and when it appeared at 1.5 s (after the final Target syllable). In the FOCUS trials, the magnitude of the N1 evoked by the Supertarget was not significantly different when the delay was 1 s and when it was 1.5 s for either Spatial Attention ($t_{(11)}$=1.61, p=0.13) or Talker Attention ($t_{(11)}$=0.31, p=0.76; see Figure 14c & 14g). However, in the DIVIDE trials, the N1 amplitude was larger when the Supertarget started at 1 s than when it started at 1.5 s, for both Spatial Attention ($t_{(11)}$=2.31, p=0.041) and Talker Attention ($t_{(11)}$=2.67, p=0.022; see Figure 14d & 14h). The difference in the magnitude of the N1 for different Supertarget delays was not significantly different in the Spatial Attention and the Talker Attention conditions ($t_{(11)}$=0.07, p=0.95). Taken together, these results show that the Supertarget N1 is larger when listeners are actively listening for it and it overlaps with the ongoing Target than either when they are ignoring it or when it starts after the Target (whether or not they are ignoring it).

3.2.2 Effects of attentional state on alpha power in parietal EEG channels

Figure 15 plots the normalized time courses of alpha oscillations for both Spatial Attention (top row) and Talker Attention (bottom row), separately for EEG channels P5 (left plots, over left parietal cortex) and P6 (right plots, over right parietal cortex). In the plots, baseline is from -2.9 s to -2 s (gray area in the plots). The visual cue instructing listeners on how to direct attention in the upcoming trial occurs at -2 s. Zero corresponds to the start of the auditory stimuli, and 1 s is the first moment where the Supertarget may appear (if it appears).

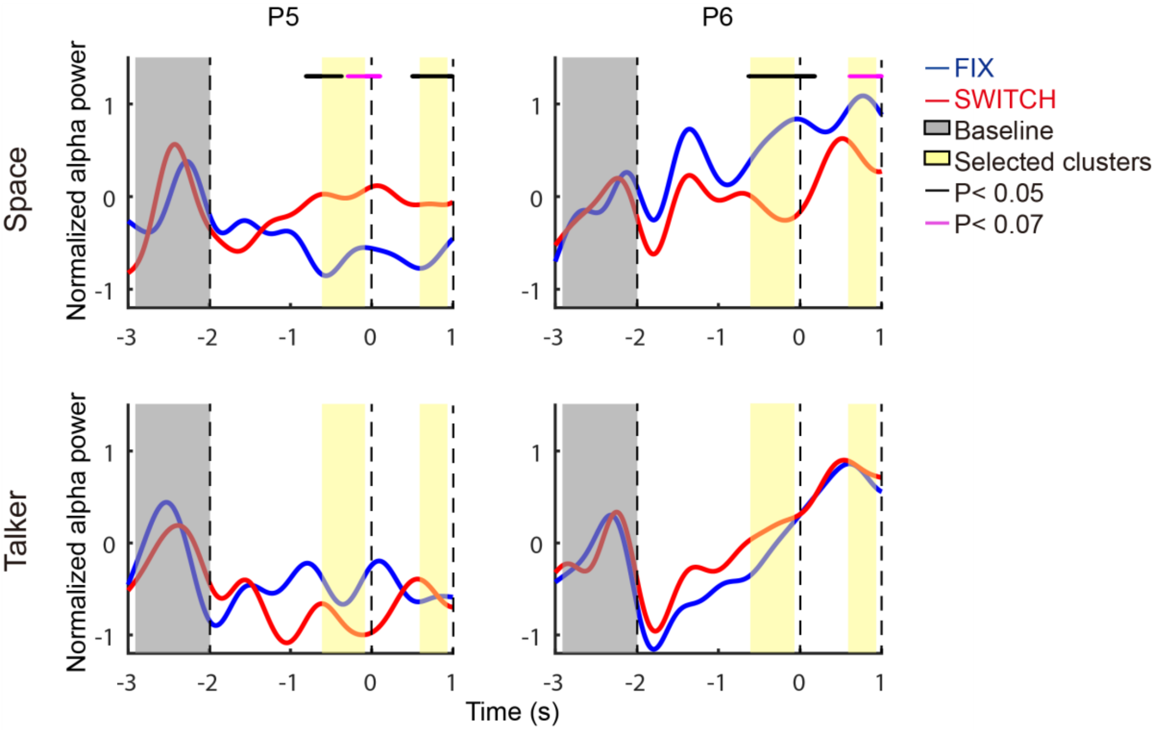

Figure 15. Alpha time course from -3 s to 1 s of the trial observed in channel P5 and P6 under FOCUS and DIVIDE strategy for Space and Pitch attention. Black bar indicates clusters of statistical significance (p<0.05). Pink bar indicates clusters of trending statistical significance (p<0.07).

In general, across both Spatial Attention and Talker Attention conditions, there is asymmetry in the alpha distribution. In the left parietal channel (P5), alpha decreases around the time of the visual cue and then remains low or rises modestly. In contrast, in the right parietal channel (P6), alpha decreases around the time of the visual cue but then rises throughout the course of the trial.

Results of clustering analysis found that in the Talker Attention condition, there were no significant temporal clusters where alpha power differed in the FOCUS and DIVIDE trials. However, during Spatial Attention, differences in alpha power for FOCUS vs. DIVIDE trials produced two temporal clusters that were either statistically significant or trending significant in both left and right parietal electrodes. (Note that a positive cluster means DIVIDE > FOCUS while a negative cluster means FOCUS > DIVIDE.) These two time periods (highlighted in yellow in Figure 15) were 1) after the cue and before the Target (in the time window from -2s to 0s, which we will call the Post-Cue Period), and 2) after the Target and before the Supertarget (in the time window from 0s to 1 s, which we will call the Post-Target Period). In both of these time periods during Spatial Attention, alpha activity over left parietal cortex was higher in the DIVIDE trials than the FOCUS trials (Post-Cue: -809 ms to -366 ms, p=0.033; -290 ms to 105 ms, p=0.061; Post-Target: 501 ms to 993 ms, p=0.036). Conversely, alpha activity over right parietal cortex was significantly lower during the DIVIDE trials than during the FOCUS trials in the Post-Target Period, and there was a trend in this direction in the Post-Target Period (Post-Cue: -621 ms to 182 ms, p=0.025; Post-Target: 606 ms to 1083 ms, p=0.066).

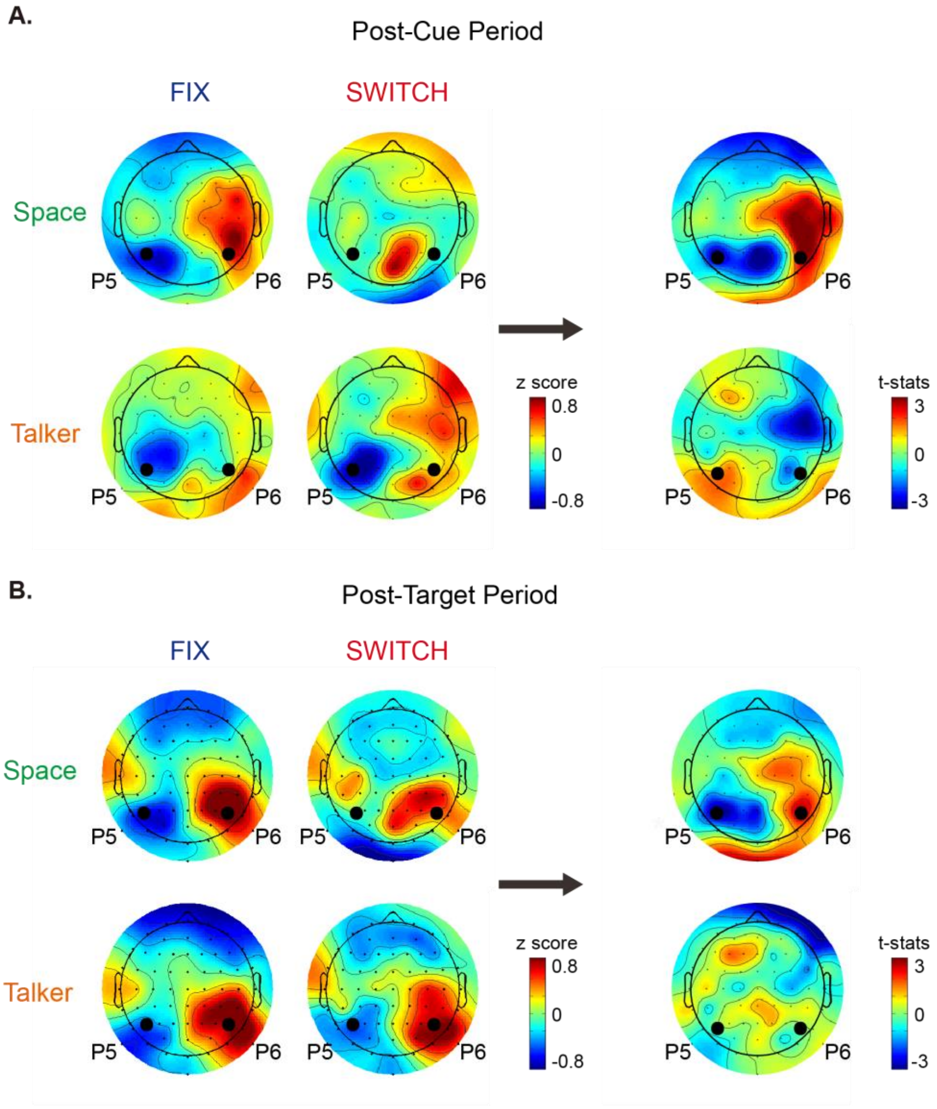

Figure 16. A) Topography plots of alpha power within selected cluster during the Post-Cue period (-621 ms – -50 ms). B) Topography plots of alpha power within selected cluster during the Post-Target period (606 ms – 950 ms). Average of alpha spatial z-score is shown for each of the four attention strategies. T-stats of alpha power comparison between FOCUS and DIVIDE condition is shown in the third column for both Space and Pitch attention. Black dots indicate the channel of interest (P5 and P6). Black dots represent channels of interest (P5 and P6).

To get a more complete picture of the differences in listening strategy during Spatial Attention and Talker Attention, we compared the topographies of alpha power for FOCUS and DIVIDE trials during both the Post-Cue Period and the Post-Target Period (see Figure 16A & 16B). For this visualization, we combined the adjacent significant and trending significant temporal clusters common to both left and right channels and dropped out the portions close to the responses evoked by Target or Supertarget onsets, resulting in one cluster during the Post-Cue Period (-621 ms to -50 ms) and one during the Post-Target Period (606 ms to 950 ms). The resulting time periods are temporally far from and should not be strongly affected by evoked activity in response to onsets of the visual cue, the Target, or the Supertarget, but were after subjects were aware of the kind of trial they were about to perform.

The Post-Cue and Post-Target cluster show similar scalp distributions of alpha power. In the FOCUS trials of the Spatial Attention condition, alpha power decreases in the left parietal channels (top row, left plots in Figures 16A and 16B) and increases strongly in right parietal channels (contralateral to the possible direction of Supertarget). In contrast, in the DIVIDE trials of the Spatial Attention condition, the alpha power is less strongly lateralized (top row, middle plots). In Talker Attention trials, alpha power decreases weakly in left parietal channels (bottom row, left plots) and increases weakly in right parietal channels (bottom row, middle plots) in both FOCUS and SWITCH trials. To analyze the difference in alpha power at each sensor between FOCUS and DIVIDE trials, we did a paired t-test for each channel and plotted the t value of each channel on the scalp topography (Figure 16A & 16B right column). During Spatial Attention, this analysis shows that alpha decreases in the left parietal sensors and increases in the right parietal sensors in both clusters (top row, right plots) ; however, there is almost no difference in alpha power distribution during Talker Attention in either the Post-Cue or the Post-Target cluster (bottom row, right plots).

4 Discussion

4.1 Spatial Attention performance is less robust than Talker Attention performance

Our behavioral results show that dividing attention to monitor multiple streams is cognitively more costly in Spatial Attention than in Talker Attention. Specifically, on trials where no Supertarget appears, listeners perform worse when they have to monitor for a potential Supertarget compared to when they can fix their attention to the original Target. Given that these two types of trials present physically identical stimuli, the observed “monitoring cost” in performance is due to differences in the top-down listening strategy that the listeners employ when focusing attention versus when monitoring for a second stream (Figure 12). This finding is consistent with past results that spatial features are weaker cues for grouping than spectro-temporal features on the syllabic level (Culling and Stone, 2017; Darwin et al., 2003; Helfer and Freyman, 2005; Noble and Perrett, 2002; Schwartz et al., 2012). When spatial attention is divided so the attention spotlight covers a wider area rather than a specific location, spatial differences within the attentional spotlight no longer benefits selective attention and spatial release from masking (SRM) (Freyman et al., 2001; Kidd et al., 1998). In this case, the Target and Supertarget within the divided attention spotlight would not be separated on the syllabic level based on their spatial distinctions and thus grouped into one single stream. The increased difficulty in streaming within divided spatial attention explains the increased “monitoring cost”. Despite being a weak cue for grouping at syllabic level, when target and maskers have distinct spatial locations, spatial features are strong cues for selective attention and help separate multiple streams (Maddox and Shinn-Cunningham, 2012) to achieve SRM. Our findings showed that in FOCUS trials during spatial attention, single differential features condition (only spatial difference) and double differential features condition (both spatial and voice difference) have comparable performance (Figure 11).

Compared to Talker Attention, our behavioral results demonstrated that Spatial Attention is harder to maintain and more susceptible to disruption by bottom-up interruptions (the appearance of the Supertarget). In FOCUS trials with an interfering Supertarget, Spatial Attention but not Talker Attention was affected by the appearance of the Supertarget, reflecting a bottom-up interruption in processing caused by the appearance of a new stream (Figure 13). It is notable that the appearance of the Supertarget not only affected recalling the final Target syllable, which overlapped with Supertarget, but also affected recalling the initial syllable, which did not overlap with Supertarget. This result demonstrates that the appearance of the Supertarget interferes with working memory processes: not only does the interruption interfere with the processing of the subsequent Target syllable, but it also affects recall of the already heard Target syllables. Similar to the cost of dividing attention, the cost of disruption affects Spatial Attention more than Talker Attention.

Together, these results hint that behaviorally, Spatial Attention requires more cognitive resource to divide and maintain and is therefore less robust than Talker

Attention. In line with explanation suggested by previous studies (for review, see Shinn-Cunningham et al., 2017), processing spatial cues is likely to involve more cognitive computation steps with binaural inputs comparison and is therefore less automatic than the processing of spectro-temporal non-spatial features. Lower reliability of spatial cues may be another reason (Shinn-Cunningham et al., 2017). In natural settings, reverberation and interference may further complicate the process of attending to spatial features and make it a less effective attentional strategy. Together, it is implied that spatial attention requires deliberate top-down effort to maintain. As a consequence, attending spatially is easily disrupted, and dividing attention to monitor two locations impairs the ability to process the attended information.

Moreover, we observed that Spatial Attention builds up over time, as evidenced by the fact that the cost of dividing attention increases from syllable to syllable, reflecting a buildup of attention (e.g., see Dai et al., 2018). This pattern is consistent with past results showing that when listeners sustain Spatial Attention, they become better and better at filtering out distracting sounds from neighboring directions (e.g., Best et al., 2008). Because of this sluggish buildup process, the cost in behavioral performance is larger for the final Target syllable in Spatial Attention but not in Talker Attention.

4.2 Dividing spatial and non-spatial auditory attention evoke different patterns of alpha activity

Our EEG results demonstrate that dividing Spatial Attention, but not Talker Attention, modulates preparatory alpha activity over parietal regions. We suggest that Spatial Attention would engage a fronto-parietal multisensory attention network that generates and modulate topographic specific alpha activities, but that Talker Attention would not rely on this same network.

Consistent with previous studies, our results of anticipatory alpha activities during spatial attention shows a lateralization pattern where alpha power increases in the contralateral hemisphere of unattended stimuli and decreases in the ipsilateral hemisphere of unattended stimuli (Atilgan, 1998; Kelly, 2006b; Worden et al., 2000a). Specifically in our findings, in the FOCUS trials during both Spatial Attention and Talker Attention blocks, we observed alpha to be enhanced over the right parietal cortex, as listeners knew that any to-be-ignored Supertarget would appear on the left. In Spatial Attention under DIVIDE condition, this lateralization pattern disappears as the potential left Supertarget is no longer an unattended stimulus, but another stream the listeners actively monitor. The modulation of parietal alpha power based on attended spatial locations supports the hypothesis that parietal alpha generated by the fronto-parietal spatial attention network plays an active role in the inhibitory control of unwanted stimuli (F Dosenbach et al., 2007; Foxe et al., 1998; Klimesch et al., 2007b). Consistent with our observations, (Klatt et al., 2018) showed that there is no posterior alpha lateralization during an auditory detection task where attention is not focused on a specific location.

However, this modulation of alpha was not observed in the DIVIDE condition of Talker Attention. The lack of parietal alpha modulation by Talker attention could likely support the theory that “what” pathway does not utilize the same fronto-parietal network as the “where” ventral pathway. This is in line with results from a meta-analysis on previous fMRI studies that shows while spatial auditory tasks ubiquitously involve inferior parietal lobe (IPL), less than half of the non-spatial studies reported IPL activation (Arnott et al., 2004). Apart from parietal region, we did no observe alpha modulation by Talker attention in any other cortical areas. Even though there is neurophysiological evidence that pitch selectivity exists in auditory cortex (Bendor and Wang, 2005; Plack and Barker, 2014)doubts remain about whether alpha oscillation can be observed in auditory cortex by EEG (Frey et al., 2014; Weisz et al., 2011). The absence of attentional modulation in the alpha activities in our Talker Attention could be due to limitations of EEG in measuring deep sulcus region like the auditory cortex. Alternatively, our findings could add on to the evidence that alpha oscillation during spatial attention does not play the same functional roles in object-based attention, which is consistent with a past study (Slagter et al., 2016) that shows when maintaining visuospatial attention, attending to a certain object in the same location does not modulate alpha power.

4.3 Induced alpha oscillation and evoked N1 response reflect different ways of deploying selective attention

We found that attentional modulation of evoked N1 response was statistically similar between Space and Talker Attention. In both conditions, the N1 response to attended Supertarget is stronger than unattended Supertarget (Figure 14). The strength of the attentional modulation was similar. The attentional N1 modulation we observed is consistent with previous ERP studies on auditory spatial attention modulation (Hillyard et al., 1973; Teder-Sälejärvi and Hillyard, 1998) and feature-based attention modulation (Alho et al., 1987; Woldorff et al., 1990). Notably, this modulation of Supertarget evoked response was only observed when the Supertarget overlapped with the Target (Supertarget a 1 s) but not when they were present after the Target (Supertarget 1.5 s). The presence of attentional modulation only when competing streams are present indicates that this modulation functions as a separation mechanisms of attended and unattended streams, in line with the hypothesis that N1 represents as sensory gain control that benefits the orientation of attention to an attended stimuli (Luck et al., 1990).

Further comparison between different timings revealed that when competing stream was present, N1 evoked by the attended Supertarget is enhanced, but N1 response to the unattended Supertarget is not suppressed (Figure 14). A number of past studies have reported attentional modulation of ERP components, either through suppression of distractor stimuli processing (Couperus, 2009; Handy et al., 2001), facilitation of attended stimuli (Barnhardt et al., 2008; Luck et al., 1994; H. A. Slagter et al., 2016), or both (J.W. Couperus and Mangun, 2010; Gazzaley et al., 2005). However, it remains unclear whether the top-down influence should be interpreted.

Taken together, our findings show that spatial and non-spatial auditory attention modulate evoked N1 responses similarly but alpha oscillation differently. While induced alpha was traditionally interpreted as a marker of information-gating processes, our results on evoked responses suggest that N1 modulation could be a result of an enhancement mechanism to increase sensory input gain of attended stimuli. Therefore, our results provide evidence for the theory that induced alpha activities and evoked sensory responses reflect different attentional strategies at different attentional processing stages. While preparatory alpha plays a functional role of suppressing task-irrelevant cortical processes as early as anticipatory period (Jensen and Mazaheri, 2010b; Klimesch, 2012b), early cortical evoked responses (e.g., N1) are modulated by neural processes to aid segregating and selecting relevant stimuli rather than suppressing irrelevant stimuli (H. A. Slagter et al., 2016).

Conclusion

The present study compares behavioral performance and the neural dynamics of auditory spatial and non-spatial attention. With human EEG and a divided attention task that allows direct comparison of these two forms of attention, we demonstrate two main findings. Behaviorally, the “where” and “what” attentional strategies influence attentional performance. Neurally, these two strategies modulate evoked response similarly, but modulate preparatory alpha oscillation differently. Based on our results, we propose that spatial and non-spatial attention engage different cortical regions and processes, which results in different behavioral performance in challenging conditions that require divided attention.

CHAPTER THREE

Causal links between parietal alpha activity and auditory spatial but not non-spatial auditory attention

Yuqi Deng[1], Robert Reinhart[2], Inyong Choi[3], Barbara Shinn-Cunningham[1,4]

[1] Department of Biomedical Engineering, Boston University, 610 Commonwealth Ave., Boston, MA, USA, 02215;

[2] Department of Psychological & Brain Sciences, Center for Systems Neuroscience, Center for Research in Sensory Communications and Neural Technology, Boston University, Boston, MA 02215, USA

[3] Department of Communication Sciences and Disorders, University of Iowa, Iowa City, IA, 52242;

[4] Carnegie Mellon Neuroscience Institute, Carnegie Mellon University, Pittsburgh, PA 15289;

Abstract

In vision, spatial attention relies on retinotopic maps of space in parietal cortex, which are dominated by a representation of contralateral exocentric space. Visual spatial attention towards a particular exocentric location in space leads to strong alpha (8-12 Hz) oscillations over ipsilateral parietal cortex; these alpha oscillations are thought to represent an “idling” state associated with suppression of information in the corresponding parietal map representing ignored locations. Recent reports suggest that these parietal maps are engaged not only by visual spatial attention, but also by spatial auditory attention; however, there is no causal evidence implicating parietal maps of space in control of attention to sound from a particular direction. Here we performed high definition transcranial alternating current stimulation (HD-tACS) to interfere with parietal activity in right intraparietal sulcus (rIPS), which should contain the only strong representation of exocentric space to the left of listeners. Stimulation was administered on human subjects who were asked to direct either spatial (attend left or attend right) or non-spatial (attend male or attend female talker) attention when presented with competing streams of spoken syllables. We found that alpha (10Hz) HD-tACS of right parietal cortex interfered with attention directed towards left auditory space; however, this stimulation had no effect on attention to right auditory space. The stimulation had no effect when attention was directed to non-spatial sound features, when parietal maps should play no role, and had no effect for stimulus conditions where spatial attention was already disrupted by bottom-up discontinuities of talker gender in the syllables in the attended direction. Further, theta (6Hz) stimulation did not have any stimulation effect.

Our findings demonstrate a frequency specific and hemispheric specific causal link between lateralized parietal activity and active spatial information gating of auditory inputs.

1 Introduction

Parietal cortex is known to interact with regions in frontal cortex in controlling visual spatial attention. In particular, a series of retinopically mapped representations of visual inputs, which ascend from visual cortex up the intraparietal sulcus (IPS), form a key part of a fronto-parietal spatial attention network long implicated in control of visual spatial attention. Although less well studied, a few recent studies suggest that the same fronto-parietal network is engaged by auditory spatial processing, including auditory spatial attention.

Retinotopic maps in IPS represents sensory inputs from exocentric space in a lateralized, but asymmetric, code. Specifically, information about inputs in contralateral exocentric locations dominate representations in IPS. However, whereas inputs in left exocentric space are only represented in contralateral right IPS, inputs in right exocentric space are represented strongly in left IPS as well as weakly in right IPS. This asymmetry helps explain how hemispheric neglect can arise with lesions of right IPS, by interfering with the sole representation of inputs to the left of an observer (e.g. see Farah et al., 1989).

When observers direct visual attention to either the left or right, the power of neural activity in alpha frequencies (8-14 Hz) tends to increase in IPS ipsilateral to the direction of spatial attention (Macaluso et al., 2002; Weisz and Obleser, 2014). Such observations have led to the idea that parietal alpha oscillations are associated with inhibition of information in the corresponding parietal representation, which is dominated by information from to-be-ignored directions (Kelly et al., 2006; Klimesch, 2012a; Worden et al., 2000b). A few studies find similar patterns of lateralized alpha activity, increasing in the parietal areas ipsilateral to the focus of attention, during auditory spatial attention (S Banerjee et al., 2011; Mehraei et al., 2018; Wöstmann et al., 2016b), but not during non-spatial attention (Mehraei et al., 2018). Pilot data from our own laboratory reveal a systematic shift of alpha lateralization as the focus of auditory spatial attention shifts from far-left to far-right exocentric locations (Deng et al, 2017; see Figure 17).

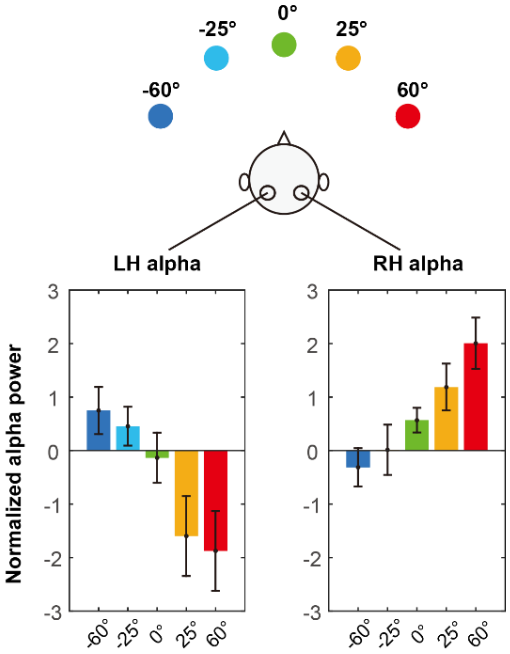

Figure 17. Alpha activities during auditory spatial attention. Spatial attention was covertly directed to five different spatial locations represented by different colors. Parietal alpha power on the left and right hemisphere during the attention preparation period was plotted on the left and right bar graphs respectively. Twenty-six subjects were tested in this study (N=26).

However, all previous evidence about the role of parietal representation in the control of auditory spatial attention has been correlative. We know of no studies that address the question of whether parietal activity arises as a byproduct of auditory spatial processing, or whether this activity is necessary to achieve desired auditory spatial attention control.

Brain stimulation methods such as transcranial magnetic stimulation (TMS) and transcranial electrical stimulation (tES) allow direct investigation into the causal role(s) of cortical regions in specific behavioral function(s) (Bestmann et al., 2015; Dayan et al., 2013). However, controversy remains about whether actual entrainment to the targeted neural oscillation is possible (Herrmann et al., 2016; Thut et al., 2011; Vossen et al., 2015). Several brain stimulation studies in the visual domain have led to conflicting results on whether the effects of externally stimulated parietal alpha is bilateral or spatially selective, which is crucial to determine whether the stimulation influence was caused by simple disruption or entrainment of oscillatory activity (Thut et al., 2011). While a TMS study by (V Romei et al., 2010) found spatially specific effects of alpha stimulation, another TMS study by (Capotosto et al., 2009) and a tACS study by (Brignani et al., 2013) observed impairment of attention performance during stimulation, but the impairment was not spatially selective. (Veniero et al., 2017) observed inconsistent effects. Apart from spatial specificity, controversy also exists in whether the stimulation effect is frequency specific. (Brignani et al., 2013) observed marginal frequency specificity, implying that the observed stimulated effect is likely not caused by entrainment. All of the aforementioned brain stimulation studies were done in the visual domain. To our knowledge no studies have been done causally linking parietal alpha oscillation to auditory spatial attention.

Methodological accounts have to be considered in these brain stimulation studies. TMS is usually administered through pulses to cause direct neural firing, which is not confined to the single stimulation frequency (Herrmann et al., 2013). Contrary to TMS, tACS modulates neuronal membrane potential rather than causing direct firing (Reato et al., 2010; Tavakoli and Yun, 2017). Due to the sinusoidal waveform, tACS has been suggested to be less likely to entrain to the unintended frequency as compared with TMS (Herrmann et al., 2013). However, traditional tACS is usually delivered with large pads (20–30cm^2) and therefore stimulates a broad area between the stimulation electrodes (Datta et al., 2009; Villamar et al., 2013). The lack of focality in traditional tACS could potentially spread the electric current across both hemispheres and complicate results.

In the present study, we used high-definition transcranial alternating current stimulation (HD-tACS) to disrupt parietal activity in order to evidence of a causal relationship between parietal alpha oscillation and auditory spatial attention. We hypothesize that stimulation effect should be spatial specific, frequency specific and task specific to reflect entrainment of oscillatory activities. To improve spatial resolution of stimulation, we applied HD-tACS to stimulate alpha in right IPS (rIPS) (Figure 21A). Recently, HD-tACS using an electrode ring configuration (e.g., an anode flanked by multiple cathodes) has been shown to create a more focused electrical current and increase anatomical targeting of a cortical region, relative to traditional tES methods (Helfrich et al., 2014; Nguyen et al., 2018) (Reinhart, 2017). In addition, improved computational models of current flow have sharpened our reasoning about where in the brain the electric current is traveling during stimulation (Datta et al., 2009; Truong et al.,

2013) (Reinhart et al., 2017). In the present study, we leveraged these methodological advances to more precisely target rIPS and examine the causal relationship between parietal cortex and auditory spatial attention.

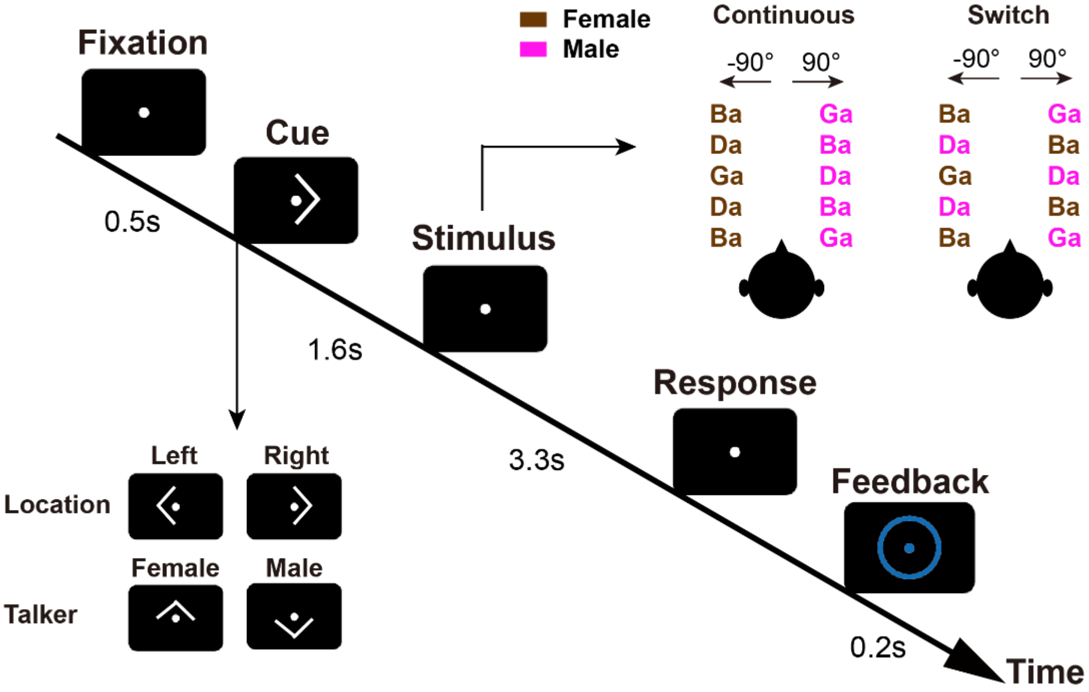

Figure 18. Illustration of task paradigm. Subjects were instructed to initially fixate their gaze on a fixation dot. A visual cue later informed them the feature of the target stream to attend to. A target and a distractor stream was presented after the visual cue. Subjects were instructed to count the number of /ga/s appeared in the 5-syllable target stream and report their response as soon as possible. The target and distractor stream could either be continuous (same talker in spatial attention; same location in talker attention) or changes the task irrelevant feature every syllable. A feedback is given to inform the subjects of the correctness of the task.

We designed a task similar to that performed by (Mehraei et al., 2018) (Figure 18), who found parietal alpha lateralization during auditory spatial attention but not non-spatial attention. Moreover, their results suggest that during spatial attention, if the current auditory object is disrupted by change in a task-irrelevant feature, alpha lateralization pattern dissipates to engage in new auditory object formation (Figure 19). Therefore, the design of this task provides an extra control for us to compare between two conditions: the continuous condition when alpha lateralization is important for maintenance of endogenous spatial attention and the switch condition when it is overridden by bottom-up object disruption.

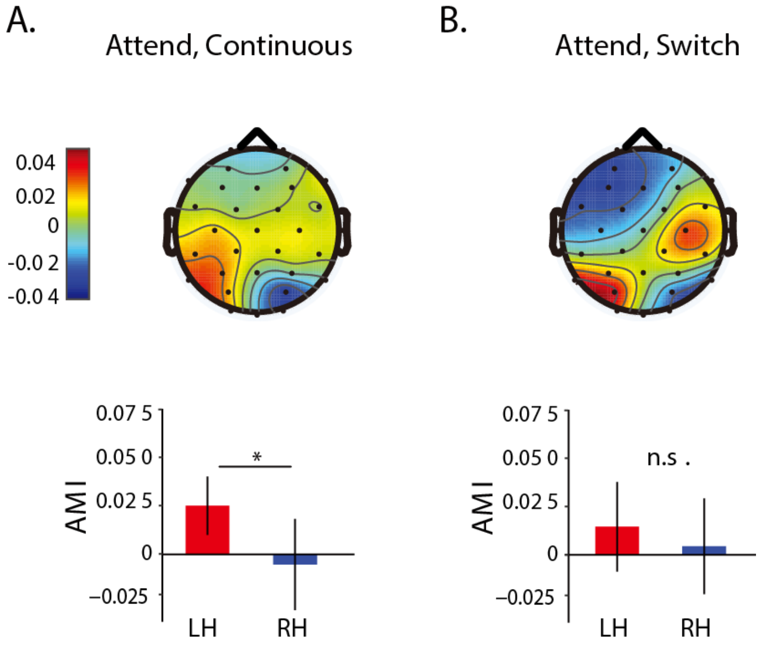

Figure 19. Adapted with permission from (Mehraei et al., 2018). A. Alpha lateralization during continuous spatial attention focused to the left. B. Alpha lateralization pattern dissipates when attended stimuli changed task-irrelevant features.

Our experiment design included two experiments. In Experiment 1, we stimulated at alpha frequency (10Hz). In Experiment 2, we stimulated at theta frequency (6Hz). We tested three hypotheses: 1) Alpha stimulation in parietal cortex affects spatial attention in a hemispheric specific manner (Figure 20). More specifically, based on the neglect model that left parietal cortex primarily represents the contralateral hemifield while right hemisphere represents both hemifields (Huang et al., 2014; Okazaki et al., 2015; Teshiba et al., 2013), we hypothesize that external alpha stimulation in rIPS would inhibit information processing from the contralateral side (left hemifield), reverse the natural alpha lateralization pattern and therefore impair behavioral performance when attention is focused on the left. On the other hand, performance when attending to the right would increase due to maintenance or reinforcement of the natural alpha lateralization pattern and inhibition of the unattended stimuli from the unattended left side. 2) IPS related alpha stimulation does not modulate performance in non-spatial attention. In addition, we hypothesize that spatial attention manipulation by alpha stimulation only exists in trials where alpha lateralization is not overridden by bottom-up disruption of streaming that leads to object reformation. 3) HD-tACS stimulation effects is frequency specific, where theta stimulation does not affect spatial attention similarly as alpha stimulation. We included Experiment 2 as a control session to ensure that the simulation effects we observed from alpha stimulation is frequency specific rather than stemming from general disturbance of the affected cortical area.

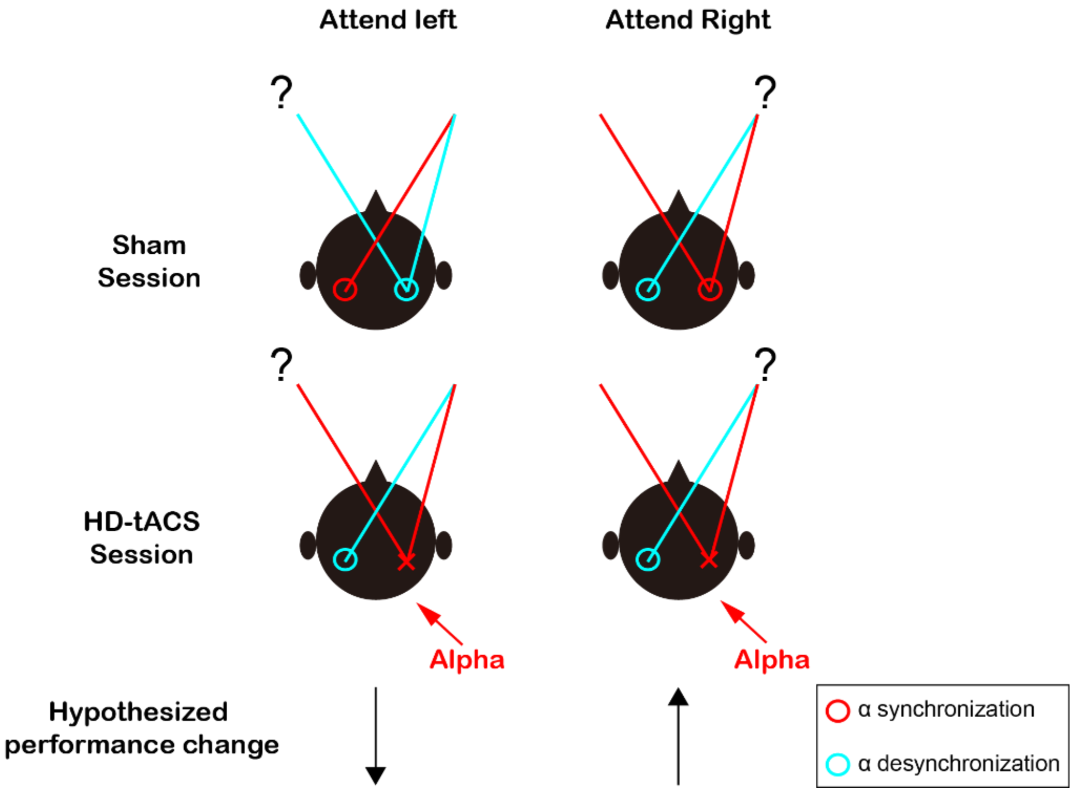

Figure 20. Illustration of hypothesis. Information from left hemifield is represented in the contralateral hemisphere, while information from the right hemifield is represented in both hemispheres. Without stimulation (top row), when attending to left, alpha power would decrease on the contralateral right hemisphere and increase on the left hemisphere. The lateralization is flipped when attending to right. With stimulation (middle row), the natural lateralization pattern is reversed when attending left and reinforced when attending right. Therefore, we predict (bottom row) that behavioral performance when attending left will be impaired while that when attending right will be enhanced.

2 Methods and Materials

2.1 Subjects

Twenty paid volunteers participated in Experiment 1. The participants were within the age of 18-24 years (M=21.15, SD=3.01; 13 females, 7 males). Eighteen paid volunteers participated in Experiment 2. The participants were within the age of 18-24 years (M=22.11, SD=2.4; 15 females, 3 males). None of the participants had audiometric thresholds greater than 20 dB for any frequency from 250 Hz to 8k Hz. All participants gave informed consent as approved by the Boston University Institutional Review Board.

2.2 Stimuli and procedure

Task:

Subjects performed an attention task in variation to that designed by (Mehraei et al., 2018)(Figure 18). At the beginning of each trial, subjects fixated their gaze on a dot presented at the center of the screen. A visual cue was presented 0.5 s after the fixation dot and lasted for 0.4 s to instruct the subjects which sound stream to attend to. The visual cue of left or right arrow indicated the location of the to-be-attended target stream (left or right); while an up or down arrow indicated the talker identity of the target stream (female or male). Sound stimuli made up of a target stream and a distractor stream started 1.2 s after the visual cue disappeared. Subjects were asked to report the number of /ga/ syllables that appeared in the target stream regardless of the feature change in the target stream. At the end of every trial, feedback was given to inform subjects whether their response was correct. The trial order was randomized and each condition had an equal probability of being presented.

During each experimental session, each subject performed 3 blocks of randomized trials. In each block there were a total of 208 trials, 52 trials in each attention condition (left, right, female or male). After every 48 trials, subjects were given a brief resting period of 10 s. Prior to the experimental session, each subject receives two training blocks. In the first training block, subjects practiced counting the number of /ga/ appearing in a 5-syllable stream without a distractor stream. Subjects practiced until their accuracy reached 80%. In the second training block, subjects performed the experiment attention task until their accuracy reached 50%, which is above the chance level of 17%.

Stimuli:

Sound stimuli consisted of consonant-vowel syllables (/ba/, /da/, & /ga/), each 388 ms in duration. These syllables were recorded from a female talker (F0 =189 Hz, estimated by Praat software) (Boersma, 2001) and a male talker (F0 =125 Hz, estimated by Praat software). Following the visual cue, two spatially separated sound streams were presented (target and distractor stream) in an interdigitated manner, with the distractor stream leading the target stream by 180ms. Both streams consisted of 5 syllables randomly chosen with replacement out of the sets of three syllables. A constraint is that any syllable in one stream could not be the same as its adjacent counterparts in the other stream. To avoid engagement of the other form of attention, in spatial attention the target stream could be spoken by either the female or the male talker; in talker attention the target stream could be from either the left or right location. Each syllable in one stream was separated by 433ms from onset to onset. Sound stimuli was spatialized to approximately ±90° azimuth (0° elevation). The spatialized sound was constructed by convolving sound stimuli with measured KEMAR (Knowles Electronics Manikin for Acoustic Research) HRTFs (Gardner & Martin, http://sound.media.mit.edu/resources/KEMAR.html). Stimuli were presented via earphones (ER-2, Etymotic Research, Inc.) in an Eckel sound-proof booth. All sound stimuli were presented at a sound pressure level of approximately 75 dB.

The task irrelevant feature in spatial attention is talker identity, and that in talker attention is spatial location. In the continuous condition, the target and distractor stream maintained its task-irrelevant feature throughout the presentation; in the switch condition, the two streams switch task irrelevant feature every syllable.

2.3 Experiment design

We conducted two Experiments. In Experiment 1, HD-tACS was conducted at alpha frequency (10Hz). In Experiment 2, HD-tACS was conducted at theta frequency (6Hz). In both experiments, each subject had two different stimulation conditions (a 1.5 mA HD-tACS session and a sham control session) administered on two separate days, order counterbalanced across subjects. In Experiment 1, the two testing days were separated by 1–14 days (M = 2.45 days). In Experiment 2, the two testing days were separated by 1–16 days (M = 4.61 days). To ensure that the testing procedure does not bias behavioral performance, subjects were blinded to the condition of the stimulation. In both HD-tACS session and sham control session, stimulation was delivered over the right parietal cortex.

Figure 21B illustrates the experimental procedure. During each testing session, subjects performed 3 blocks of the experimental task, each consisting of 208 trials and lasting approximately 20 min. During the first block, no stimulation was administered so that this pre-stimulation block served as a baseline control for each session. During the second block, we performed HD-tACS stimulation or sham stimulation for 20 min. During the third block, no stimulation was performed so that we could examine the aftereffects of stimulation. According to evidence that alpha frequency stimulation is more effective in the darkness (Kanai et al., 2008), this experiment was performed in the dark for all but one subject due to their self-reported claustrophobic.

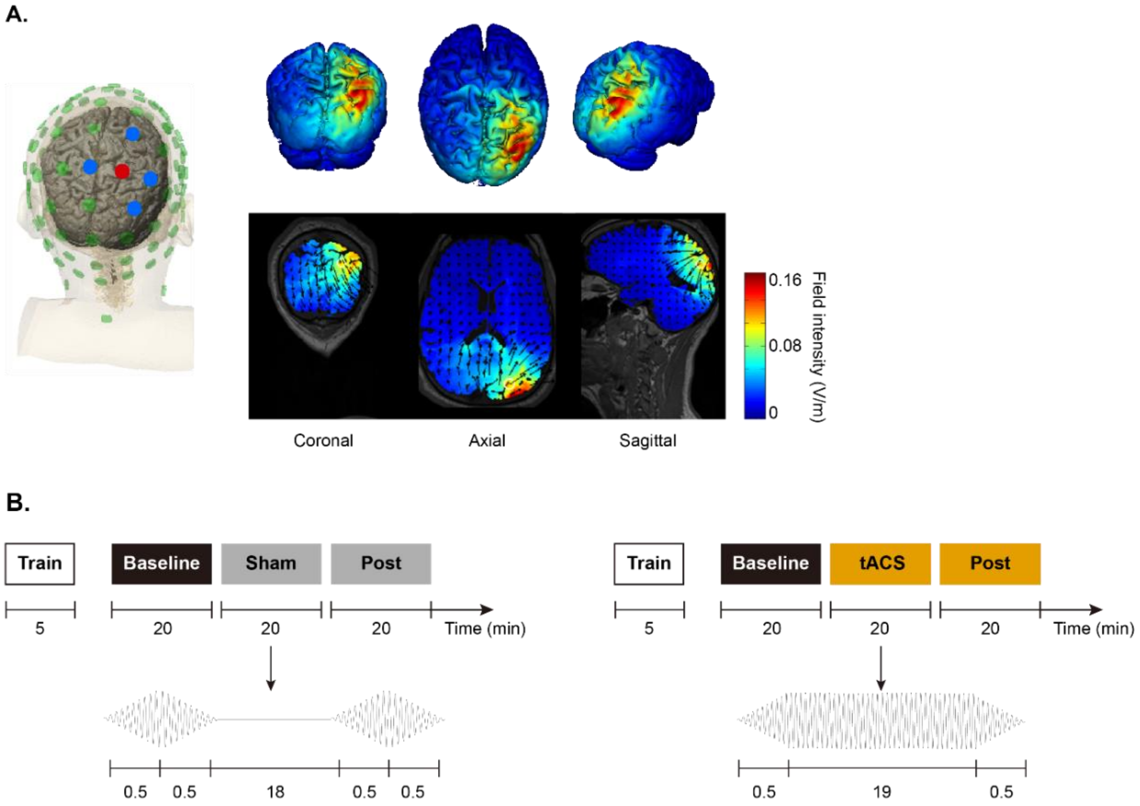

Figure 21. HD-tACS model and stimulation protocol. (A) Illustration of electrode montage and current flow model of HD-tACS. Five parietal channels (CP2, P2, P4, Pz, POz) were selected from a 64-channel EEG cap. Blue color indicates one

stimulation polarity and red indicates the opposite polarity. Together a current sink was created to target the right intraparietal sulcus (IPS). Bottom row illustrates the current flow model targeting right IPS. Top row shows the view of the 3D model from 3 different angles. Bottom row shows the 2D coronal, axial and sagittal view of a portion of IPS. White circle represents the Montreal Neurological Institute (MNI) coordinate selected (24 -70, 41). These coordinates were selected based on a previous fMRI study that mapped the location of human IPS (Swisher et al., 2007). (B) Experiment design. Stimulation was administered on two sessions for each subject: a sham session and a HD-tACS session. During both sessions, subjects performed a brief training block of the task, a pre-stimulation baseline block, a during-stimulation block and a post-stimulation block. Each experiment block lasts for approximately 20min. The waveforms of the sham and HD-tACS stimulation at 10Hz was illustrated. In the sham session, stimulation ramps up and down at the beginning and the end. While in the tACS session, stimulation ramps up and maintained the intensity until the end.

2.4 High Definition Transcranial Alternation Current Stimulation (HD-tACS)

HD-tACS was administered using the Soterix M×N-9 High Definition-Transcranial Electrical Current (HD-tES) Stimulator (Model 9002A, Soterix Medical, New York, NY). To target right IPS, an electrode montage was created based on a current flow model generated by the Soterix HD-Explore software (version 4.1, Soterix Medical, New York, NY). The major stimulating electrode was placed at P2 with a stimulation intensity of 1.5mA, and 4 return electrodes were placed at CP2 (-0.6mA), P4 (-0.225mA), Pz (-

0.075mA), and PO4 (-0.6mA). Both HD-tACS stimulation and sham session used the same electrode montage. Figure 21A depicts the electrode placement of the montage and simulated current-flow model.

In both HD-tACS stimulation session and sham control session, the waveform delivered was a bipolar sinusoidal wave at 10Hz (Figure 22B). Individual differences exist in alpha peak frequency, and we chose 10Hz for all subjects based on previous results that the mean individual alpha frequency lies between 10–11Hz (Haegens et al., 2014; Klimesch, 1997). For theta frequency we picked 6Hz which is the median peak frequency reported in theta oscillation and did not overlap with the alpha range (Jacobs, 2013). In the HD-tACS session, stimulation ramped up to 1.5mA in 30 s at the beginning of the stimulation block, and ramped down in 30 s at the end, leaving 19 min of continuous stimulation in between. During the sham session, stimulation ramped up to 1.5mA in 30 s and ramped down immediately to 0mA in the following 30 s, and similar ramping pattern was repeated again at the end of the block. During the 40 s at the beginning of stimulation in both sessions, subjects were verbally checked to ensure they are familiarized with the stimulation caused sensation and that they are comfortable to proceed with the experiment.

The stimulation electrodes were placed on HD Electrode holders (Soterix Medical, New York, NY) and fixated on a 64 channel EEG cap. The electrode holders were filled with gel to ensure impedance for each electrode did not exceed 50 k ohms prior to stimulation and below 5 k ohms during stimulation (Thair et al., 2017). The total current delivered was 1.5mA at maximum. While the most effective intensity and duration for

HD-tACS or traditional tACS stimulation was not clear, previous studies on tACS have most commonly used 1.5mA – 2mA (Helfrich et al., 2014; Klaus and Schutter, 2018). Hence we chose the intensity to be 1.5mA which is well within the safety limit (Nitsche et al., 2008).

2.5 Statistical analysis

For hypothesis testing, percentage of correct response was calculated for each attention condition (spatial attention: attend left vs right) (talker attention: attend female vs male). We then conducted baseline correction for each attention condition in the during-stimulation block and post-stimulation block by subtracting the accuracy of the baseline block from it. For pairwise comparisons of accuracy between conditions, a Wilcoxon signed rank test was performed (significance for $p<0.05$).

3 Results

3.1 Switching talker during spatial attention impairs behavioral performance

Behavioral results from both Experiment 1 and Experiment 2 replicated previous findings using a similar paradigm (Mehraei et al., 2018). We collapsed the stimulation conditions and timing of blocks. In spatial attention, subjects performed worse in the switch condition when both the target and distractor did not maintain the talker identity feature (Figure 22).

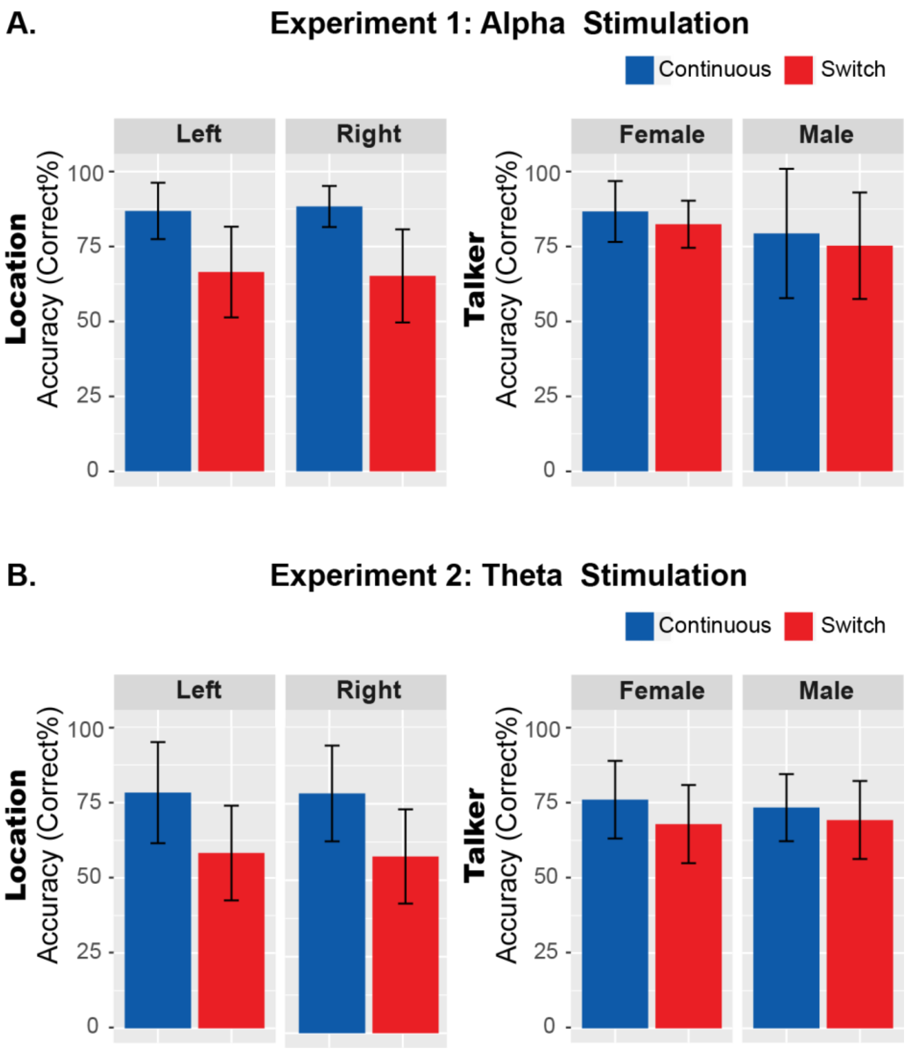

Figure 22. Behavioral performance during baseline blocks. Bar plots represents the mean accuracy collapsed across stimulation condition and timing of blocks. Left plot shows the results of spatial attention; right plot shows the results of talker attention. Error bars represent standard deviations. A. Experiment 1 with alpha stimulation. B. Experiment 2 with theta stimulation.

It is worth noting that our switch condition is different than (Mehraei et al., 2018) in that the target stream switches features more frequently (every syllable instead of every 3 syllables), and our behavioral results in the switch condition was lower than that observed in (Mehraei et al., 2018). In the continuous condition, behavioral performance of the spatial and talker attention conditions was comparable, though the performance in the attending male condition was slightly lower. For details on overall performance see Supplemental Information (5.1 Overall behavioral results).

3.2 Baseline performance

We implemented double controls for our experimental design: 1) within each session a pre-stimulation block serves as baseline control; 2) sham session was compared with HD-tACS session. The sequence of the sessions was randomized. To test whether our results could be contaminated by performance fluctuation at baseline level between test days, we averaged the accuracy scores across all attention conditions in the baseline block and compared the baseline performance between sham and stimulation sessions. After counterbalancing sequence of sham and HD-tACS session across subjects, no main effects of test day were observed for either Experiment 1 ($z(19)=0.068$, $p=0.95$, Wilcoxon rank test) or Experiment 2 ($z(17)=0.46$, $p=0.65$, Wilcoxon rank test).

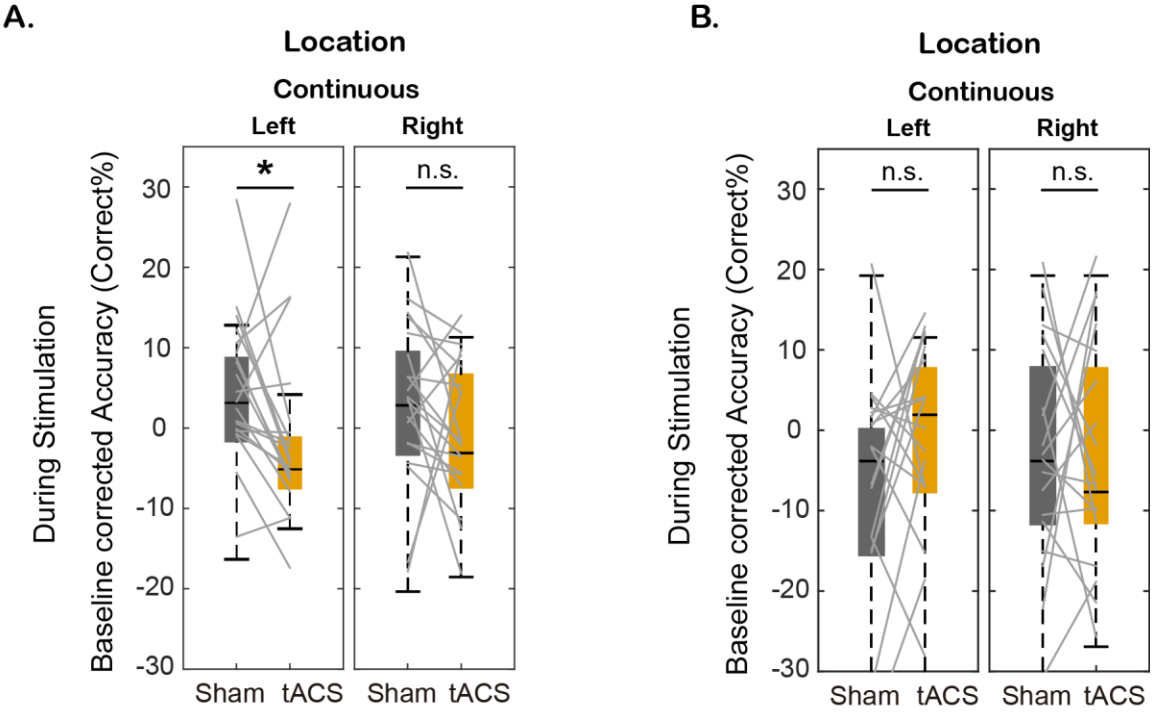

Figure 23. A. Experiment 1. Effect of alpha stimulation on behavioral performance during continuous spatial attention. B. Experiment 2. Effect of theta stimulation on behavioral performance during continuous spatial attention.

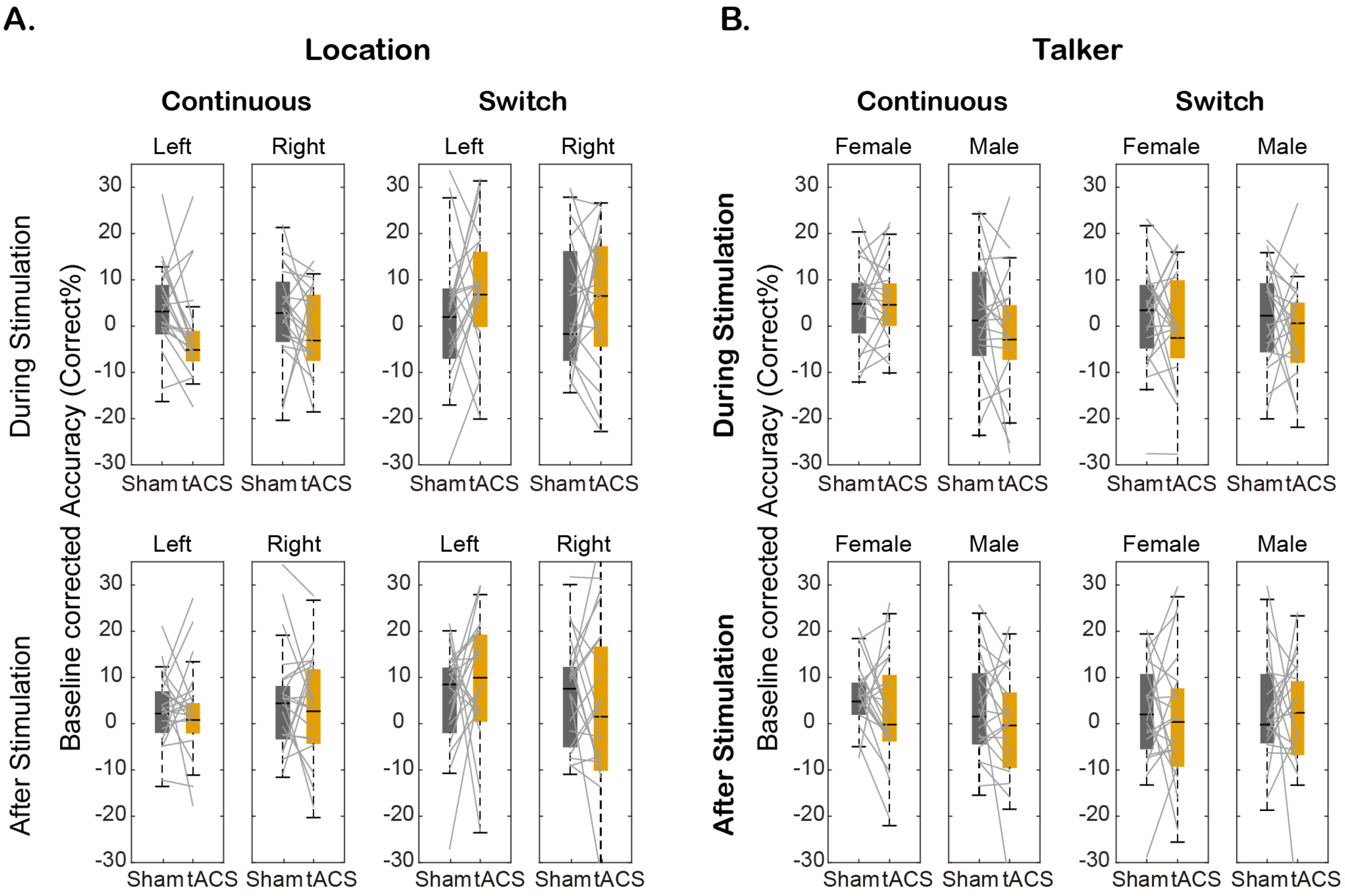
A.
Location
Continuous
Switch
Left
Right
Left
Right
During Stimulation
After Stimulation
Baseline corrected Accuracy (Correct%)
30
20
10
0
-10
-20
-30
Sham
tACS
B.
Talker
Continuous
Switch
Female
Male
Female
Male
During Stimulation
After Stimulation
Baseline corrected Accuracy (Correct%)

Figure 24. Experiment 1 results on stimulation specificity. Behavioral performance compared between sham and alpha HD-tACS sessions. A) Continuous condition under spatial attention during active stimulation. Left and right plot shows trials where attention is oriented to left and right respectively. B) Continuous condition under talker attention during active stimulation. Left and right plot shows trials where target appeared from left and right respectively. C) Comparisons of all conditions during spatial attention. Red rectangle and enlarged plot highlights the comparisons hypothesized to show significant difference between sham and tACS conditions. D) Comparisons of all conditions during spatial attention. Gray lines connect the same individual subject for comparison between conditions. *$p<0.05$.

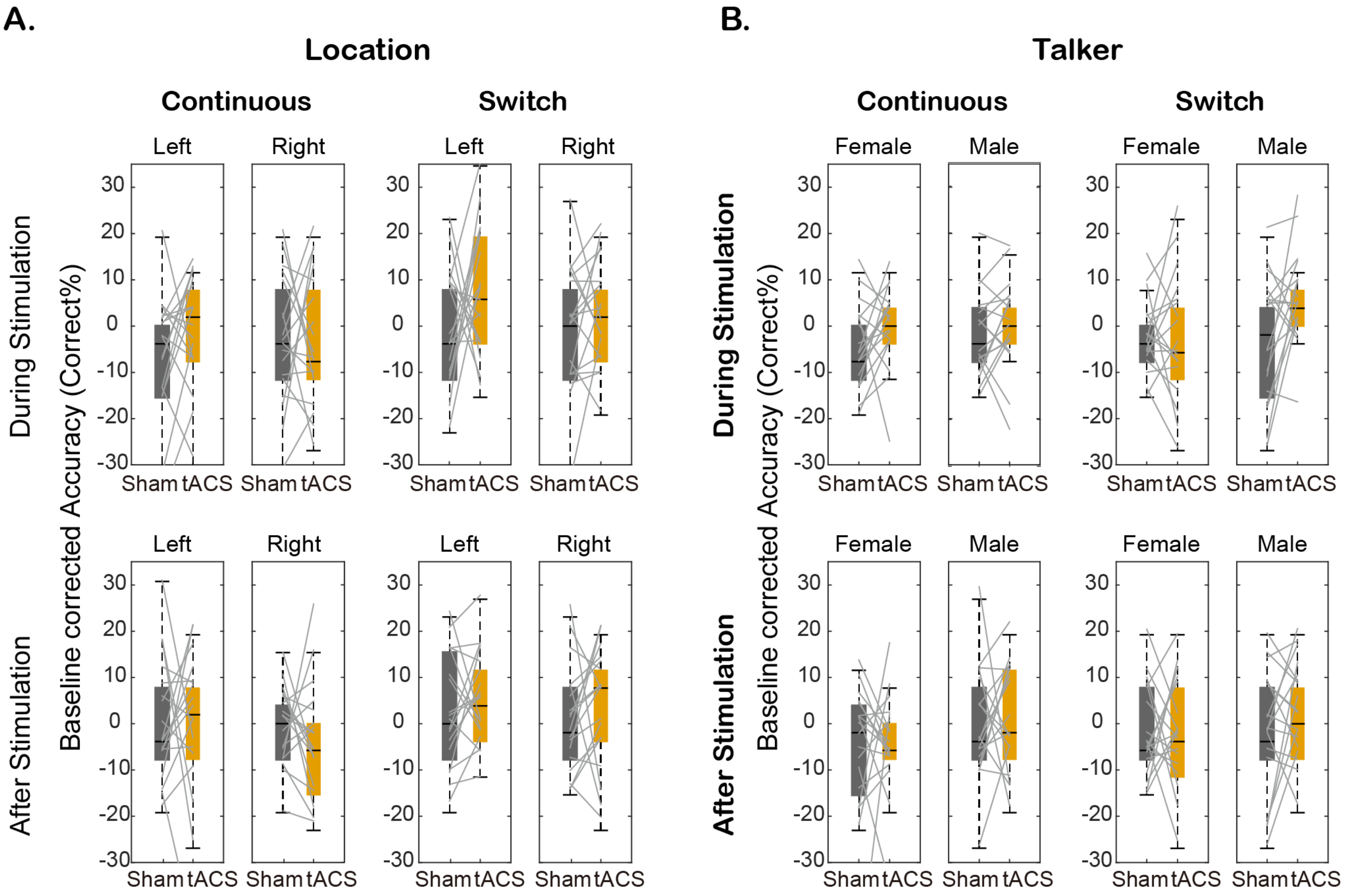
A.
Location
Continuous
Switch
Left
Right
During Stimulation
After Stimulation
Baseline corrected Accuracy (Correct%)
30
20
10
0
-10
-20
-30
Sham
tACS
B.
Talker
Female
Male

Figure 25. Experiment 2 results on stimulation specificity. Behavioral performance compared between sham and theta HD-tACS sessions. A) Continuous condition under spatial attention during active stimulation. Left and right plot shows trials where attention is oriented to left and right respectively. B) Continuous condition under talker attention during active stimulation. Left and right plot shows trials where target appeared from left and right respectively. C) Comparisons of all conditions during spatial attention. Red rectangle and enlarged plot highlights the comparisons hypothesized to show significant difference between sham and tACS conditions. D) Comparisons of all conditions during spatial attention. Gray lines connect the same individual subject for 8comparison between conditions.

3.3 Alpha but not theta HD-tACS in parietal cortex disrupts auditory spatial attention

We compared baselined corrected accuracy during sham and alpha stimulation in our test conditions: when continuous spatial attention is focused on the left and right hemifield. Consistent with our hypothesis, we found that behavioral performance was lower during HD-tACS alpha stimulation than sham session when attention is focused on the left ($z_{(19)}$=2.47, p=0.013, Wilcoxon rank test), which is the contralateral hemifield of the stimulated hemisphere (Figure 23A). However, no behavioral difference was found between sham and HD-tACS session during stimulation when attention was focused on the right ($z_{(19)}$=1.01, p=0.31), which is the ipsilateral hemifield of the stimulated hemisphere.

In contrast, in Experiment 2, theta HD-tACS did not affect spatial attentional performance. When stimulated by theta oscillation, during the stimulation block behavioral performance in the HD-tACS session did not differ from sham session whether attention is focused on the left ($z_{(17)}$=1.5, p=0.13) or right ($z_{(17)}$=0.21, p=0.84) (Figure 23B).

3.4 Alpha HD-tACS in parietal cortex is task specific

We further examined the specificity of alpha stimulation effect by comparing stimulation and sham sessions in the control conditions. Effect of alpha stimulation was only found in the continuous condition when alpha lateralization pattern was not disrupted by bottom-up feature switching (Figure 24A). In the switch condition, no effect of modulation was found either in attending left condition ($z_{(19)}$=1.21, p=0.23) or attending right condition ($z_{(19)}$=0.58, p=0.56).

We investigated in timing effect of alpha HD-tACS by comparing behavioral performance between sham and HD-tACS sessions after stimulation had finished (Figure 23A). In spatial attention, no significant modulation effect was observed in the post-stimulation block in either attending left (Continuous: $z_{(19)}$=0.87, p=0.39; Switch: $z_{(19)}$=1.03, p=0.30) or attending right condition (Continuous: $z_{(19)}$=0.23, p=0.82; Switch: $z_{(19)}$=0.66, p=0.51).

Our results showed no significant difference between sham and HD-tACS sessions in talker attention during stimulation (Figure 24B), either when attending to female (Continuous: $z_{(19)}$=0.11, p=0.91; Switch: $z_{(19)}$=0.69, p=0.49) or attending to male (Continuous: $z_{(19)}$=1.08, p=0.28; Switch: $z_{(19)}$=0.71, p=0.48). No effect of modulation was observed after stimulation when attending to either female (Continuous: $z_{(19)}$=1.02, p=0.31; Switch: $z_{(19)}$=0.33, p=0.74) or male (Continuous: $z_{(19)}$=0.98, p=0.33; Switch: $z_{(19)}$=0.095, p=0.92).

Further, we took continuous trials of the talker attention condition during stimulation and grouped these trials based on where the target was coming from, regardless of the voice to which subjects were directing attention (Figure 24B). These trials have exactly the same physical stimulus as the continuous trials in the spatial attention condition, except that subjects focus on the voice feature, not spatial location. We observed no effect of stimulation when the target is coming from either the left ($z_{(19)}$=0.23, p=0.82) or the right ($z_{(19)}$=1.30, p=0.19). Our results suggest that the alpha stimulation effect is specific to attention strategy.

3.5 Theta HD-tACS in parietal cortex does not disrupts auditory spatial or non-spatial attention

Our hypothesis focused on the differential effect between alpha stimulation and theta stimulation. In addition to our main hypothesis, we explored the effect of theta stimulation effect on accuracy in the control conditions and did not find any significant statistical effect in either spatial attention or non-spatial attention ($z<1.67$, $p>0.095$) (Figure 25).

4 Discussion

By combining HD-tACS over rIPS with an auditory attention paradigm, we demonstrated a causal relationship between parietal cortex function and auditory spatial attention. It is important to highlight the rigorous nature of the experimental design. This study has two experiments, with the 6Hz theta stimulation serves as a control session. On top of that, this study was within-subjects, sham-controlled, and employed both pre- and post-stimulation blocks on each test day to remove any differences in baseline performance that could arise across test days. This triple control design is unique in the field and represents an extremely powerful approach of gaining insight into causal relations between brain, behavior, and cognition.

Our results showed a spatial specific alpha stimulation effect. Consistent with previous visual studies using TMS for stimulation (Capotosto et al., 2009; V Romei et al., 2010), we found decreased accuracy when subjects were engaging auditory spatial attention in the hemifield contralateral to the site of stimulation and no change in behavioral performance in the ipsilateral hemifield to the stimulated site. Previous studies have observed inconsistent effect of stimulation when attending to the ipsilateral hemifield of the stimulated site. While (Capotosto et al., 2009) and (Brignani et al., 2013) found bilateral impairment in both hemispheres, (V Romei et al., 2010) found enhancement in ipsilateral hemisphere. Although one could expect that if entrained, stimulated alpha would reinforce the lateralization pattern of natural alpha, the underlying balance between inhibition and enhancement could be more complicated. A possible contribution to this inconsistency is the hemispheric asymmetry in spatial attention processing. While sensory inputs from the left hemifield are primarily regulated by the right hemisphere, inputs from the right hemifield are processed by both hemispheres. When attention is focused to the right, on the one hand rIPS stimulation could increase inhibition to the unattended side and increase performance, on the other hand information coming from the right hemifield could be inhibited too at the same time. This double representation complicates reasoning and could potentially create inconsistency in the direction of stimulation influence.

In addition, the stimulation effect we observed is specific with stimulation frequency. We conducted an additional experiment with stimulation at 6Hz, which is theta frequency closely neighboring alpha frequency (10Hz). With theta stimulation we did not observe any stimulation effects.

In addition, the alpha stimulation effects were specific to task. The comparison between continuous condition and switch condition suggests that alpha stimulation in rIPS only have behavioral consequences when task engagement requires functional alpha synchronization and desynchronization. Behavioral consequences of stimulation were observed only in the continuous condition where alpha lateralization pattern forms (Mehraei et al., 2018), but in the switch trials when alpha lateralization pattern dissipates this behavioral modulation was not found. A number of previous studies on tES have concluded that the effectiveness of stimulation depends on brain states (Nguyen et al., 2018; Reato et al., 2013). More specifically, active networks have been shown to be more sensitive to tES stimulation than inactive networks (Bikson et al., 2013; Reato et al., 2010). In line with these observations, our findings suggest that alpha modulation in rIPS is only effective when the spatial network is actively engaged during task performance, but not when endogenous changes temporarily disrupt its engagement. Furthermore, the alpha stimulation effects were task specific. We demonstrated that when listeners were performing the same physical tasks under different attention strategies, rIPS alpha stimulation influenced performance during spatial attention but not non-spatial auditory attention (Figure 23A&B). In a previous study comparing the neural mechanisms underlying spatial and non-spatial attention, we found modulation of parietal alpha during auditory spatial attention but not talker-based attention. Lateralized alpha oscillation generated by IPS could be an information gating mechanism for spatial attention only and is not involved during non-spatial attention. An alternative possible explanation is that IPS is not primarily engaged during non-spatial attention, according to a dual pathway model that in the auditory domain spatial and non-spatial information processing may engage “where” and “what” pathways respectively and employ different control mechanisms (Arnott et al., 2004; Rauschecker and Tian, 2000). While these two forms of attention are not completely isolated, spatial attention is more frequently observed to involve parietal cortex while non-spatial attention primarily engages temporal cortex and inferior frontal cortex (Arnott et al., 2004). Hence, a potential direction for futures brain stimulation studies is to target other cortical areas on the dual pathways to gain insight into the functional roles of different cortical areas during different forms of attention. However, it is still precarious to claim that our stimulation effects reflect entrainment of alpha oscillation. Critical evidence of alpha entrainment requires the stimulation effects to be specific to the hypothesized cortical process perturbed but not generally applied to any cortical functions (Bestmann et al., 2015; Thut et al., 2011). While we demonstrated spatial specificity, stimulus specificity, and task specificity, future works is needed to compare tACS stimulation effects of other frequencies to provide evidence of frequency specificity and rule out the possibility that stimulation is disrupting cortical activities in the parietal cortex.

We demonstrated that HD-tACS could manipulate neuronal oscillation during spatial attention and lead to a behavioral consequence. Stimulation focality is not a primary concern in our study. With HD-tACS, we are able to significantly improve the spatial resolution of stimulation compared to traditional tACS. However, it is worth noting that current intensity is not uniformly distributed across the targeted rIPS (Figure 22A). While IPS has a narrow and relatively long anatomy, our stimulation current is mostly focused on lower parts of IPS1 and IPS2 (Swisher et al., 2007), which has been revealed to be involved in visuospatial attention processing (Serences and Yantis, 2006; Silver et al., 2005). It is also notable that the stimulation effect we observed during active HD-tACS stimulation did not persist after the stimulation. There have been debates on whether tACS-induced neural entrainment could persist offline beyond stimulation (Antal and Paulus, 2013). While some have observed aftereffects of tACS of at least 30min (Helfrich et al., 2014; Kanai et al., 2008; Reinhart et al., 2017; Strüber et al., 2015), most in vivo animal studies have observed a lack of neural entrainment as soon as stimulation stops (Deans et al., 2007; Ozen et al., 2010). Besides the possibility that tACS entrainment does not persist offline, another possibility is that the intensity of our stimulation is not effective enough to create any aftereffect (Moliadze et al., 2012). With very few studies to reference on the effectiveness of HD-tACS intensity, we arbitrarily chose a relatively conservative and widely used stimulation intensity of 1.5mA to prevent any adverse effect of stimulation (Matsumoto and Ugawa, 2017). The presence of tACS aftereffects has been known to be nonlinear to the stimulation intensity (Antal and Paulus, 2013; Moliadze et al., 2012) and depends on brain states (Neuling et al., 2013; Nguyen et al., 2018). While a weak stimulation current can be amplified by an active network during stimulation (Reato et al., 2013), less is known about the neural activities induced by tACS after stimulation. Future endeavors are necessary to gain a better understanding of the underlying physiological effects of tACS and therefore the relationship between stimulation effectiveness and current intensity.

5 Supplemental Information

5.1 Overall behavioral results

Experiment 1. In attending left condition, performance dropped from 81.13% to 61.29% ($z_{(19)}$=4.17, p<0.001, Wilcoxon rank test); in attending right condition, performance dropped from 80.84% to 58.65% ($z_{(19)}$=4.66, p<0.001). In talker identity based attention, although average accuracy dropped slightly for both attending to female (77.53% to 71.47%) and attending to male condition (76.85% to 73.47%), neither reached statistical significance (attend female: $z_{(19)}=1.73$, $p=0.083$; attend male: $z_{(19)}=1.41$, $p=0.16$).

Experiment 2. In attending left condition, performance dropped from 86.60% to 67.21% ($z(17)=3.53$, $p<0.001$, Wilcoxon rank test); in attending right condition, performance dropped from 88.04% to 65.87% ($z(17)=4.05$, $p<0.001$). In talker identity based attention, although average accuracy dropped slightly for both attending to female (86.03% to 82.40%) and attending to male condition (78.46% to 75.58%), neither reached statistical significance (attend female: $z(17)=1.69$, $p=0.09$; attend male: $z(17)=0.74$, $p=0.46$).

CHAPTER FOUR

Impoverished auditory cues fail to engage brain networks controlling spatial selective attention

Yuqi Deng[1], Inyong Choi[2], Barbara Shinn-Cunningham[1, 3], Robert Baumgartner[1, 4]

Statement of permission:

This manuscript has been prepared for submission to a peer-reviewed journal.

[1] Biomedical Engineering, Boston University, Boston, MA 02215, USA

[2] Communication Sciences & Disorders, University of Iowa, Iowa City, IA 52242, USA

[3] Neuroscience Institute, Carnegie Mellon University, Pittsburgh, PA 15213 USA

[4] Acoustics Research Institute, Austrian Academy of Sciences, Vienna, Austria

Abstract

Spatial selective attention enables listeners to process a signal of interest in natural settings. However, most past studies on auditory spatial attention used impoverished spatial cues: presenting competing sounds to different ears, using only interaural differences in time (ITDs) and/or intensity (IIDs), or using non-individualized head-related transfer functions (HRTFs). Here we tested the hypothesis that impoverished spatial cues impair spatial auditory attention by only weakly engaging relevant cortical networks. Eighteen normal-hearing listeners reported the content of one of two competing syllable streams simulated at roughly +30 ° and -30° azimuth. The competing streams consisted of syllables from two different-sex talkers. Spatialization was based on natural spatial cues (individualized HRTFs), individualized IIDs, or generic ITDs. We measured behavioral performance as well as electroencephalographic markers of selective attention. Behaviorally, subjects recalled target streams most accurately with natural cues. Neurally, spatial attention significantly modulated early evoked sensory response magnitudes only for natural cues, not in conditions using only ITDs or IIDs. Consistent with this, parietal oscillatory power in the alpha band (8-14 Hz; associated with filtering out distracting events from unattended directions) showed significantly less attentional modulation with isolated spatial cues than with natural cues. Our findings support the hypothesis that spatial selective attention networks are only partially engaged by impoverished spatial auditory cues. These results not only suggest that studies using unnatural spatial cues underestimate the neural effects of spatial auditory attention, they also illustrate the importance of preserving natural spatial cues in assistive listening devices to support robust attentional control.

Keywords: auditory spatial selective attention, head-related transfer functions, electroencephalography

Significance statement:

Previous studies exploring the cognitive bases of auditory spatial attention rarely used individualized, fully natural spatial cues; further, no studies have explored whether or how impoverished spatial cues affect neural metrics used to quantify attentional control. We studied the influence of spatial cue richness and naturalness on the efficacy of spatial auditory attention. Comparing neural and behavioral metrics using individualized and impoverished spatial cues, we found behavioral abilities and both evoked and oscillatory neural responses are weakened or even absent with impoverished cues. Our findings suggest that unnatural and impoverished spatial auditory cues do not fully engage cognitive attentional control networks. Past results thus may fail to fully describe the cortical mechanisms underlying auditory spatial attention in real-life situations.

1 Introduction

Spatial hearing is crucial to selectively attend to sounds of interest in everyday social settings. The remarkable ability of normal-hearing listeners to focus on a sound source within a complex acoustic scene is often referred to as “the cocktail party phenomenon,” and has a rich history (Cherry, 1953). Nevertheless, the mechanisms controlling spatial selective attention are still poorly understood. Acoustically, in everyday situations, the two ears provide the listener with a listener-specific combination of spatial cues that include interaural time and intensity differences (ITDs and IIDs, respectively), as well as spectral cues caused by acoustical filtering of the pinnae (Blauert, 1997). Together, these cues, captured by individualized head-related transfer functions (HRTFs), allow the brain to create a clear, punctate internal representation of the location of sound sources in the environment (Majdak et al., 2018; Middlebrooks, 2015).

When only isolated or impoverished spatial cues are present, auditory localization performance degrades and the natural perception of external auditory objects may even collapse into the listener’s head (Baumgartner et al., 2017; Callan et al., 2013; Hartmann and Wittenberg, 1996). Nevertheless, isolated ITDs and IIDs still create a strong sense of lateralization within the head; moreover, even such highly impoverished spatial cues can be used to achieve spatial release from speech-on-speech masking, behaviorally (Culling et al., 2004; Ellinger et al., 2017; Glyde et al., 2013; Kidd et al., 2010; Loiselle et al., 2016). The relative importance of ITDs and IIDs in spatial release from masking remains unclear, with past studies reporting conflicting results when directly comparing different binaural conditions (Ellinger et al., 2017; Glyde et al., 2013; Higgins et al., 2017; Shinn-

Cunningham et al., 2005).

Previous electroencephalography (EEG) and magnetoencephalography (MEG) studies have demonstrated that rich spatial cues in sound stimuli lead to different cortical activity compared to using isolated cues during sound localization (Callan et al., 2013; Leino et al., 2007; Palomäki et al., 2005) and auditory motion processing (Getzmann and Lewald, 2010). However, the apparently minor behavioral consequences of using unnatural, non-individualized spatial cues on spatial release from masking, combined with the ease of implementing studies with simple, non-individualized spatial cues, led to their wide usage in auditory neuroscience studies (Cusack et al., 2001; Dahmen et al., 2010; Dai et al., 2018; Itoh et al., 2000; Kong et al., 2014; Sach et al., 2000). Indeed, in the auditory neuroscience literature, many studies did not even present true binaural signals, but instead studied “spatial” attention by using dichotic signals, with one sound presented monaurally to one ear, and a competing sound presented monaurally to the other ear (Ahveninen et al., 2011; Alho et al., 1999; Wöstmann et al., 2016a). These studies implicitly assumed that because listeners were able to use impoverished spatial cues to listen to one sound from a particular (relative) direction, the cognitive networks responsible for controlling spatial attention must be engaged just as they are when listening to rich, natural spatial cues. Nonetheless, it is unclear whether and how engagement of higher-order cognitive processes such as deployment of selective attention is affected by the use of unnatural or impoverished spatial cues.

Modulation of neural signatures, such as event-related potentials (ERPs) and induced oscillatory activity, is often taken as evidence of effective attentional control (Herrmann and Knight, 2001; Siegel et al., 2012). In particular, auditory spatial attention is known to modulate early sensory ERPs in the N1 time range (processing latencies of 100 to 150 ms) (Choi et al., 2013; Röder et al., 1999), whereas modulation of P1 ERPs (50 to 100 ms) has only recently been demonstrated in a free field experiment (Giuliano et al., 2014). Induced alpha oscillation (8 to 14 Hz) has been hypothesized to function as an information gating mechanism (Klimesch et al., 2007b). During auditory spatial attention, parietal alpha power often decreases in the contralateral hemisphere of attended stimuli and/or increases in the ipsilateral hemisphere (S Banerjee et al., 2011; Lim et al., 2015; Wöstmann et al., 2016a). These neural modulations constitute objective metrics of the efficacy of attentional control.

Here, we test listeners in a selective attention paradigm with simultaneous, spatially separated talkers. We use the aforementioned EEG measures to compare both perceptual ability and the neural signatures of attentional control for simulations with impoverished vs. natural spatial cues. Eighteen subjects performed an auditory spatial attention task with two competing streams located at roughly +30 ˚ and -30˚ azimuth (Figure 26). On every trial, listeners were cued by an auditory cue to attend to either the left or right stream and report the content of the cued stream. The competing streams consisted of syllables (/ba/, /da/ or /ga/) from two different-sex talkers. Sound stimuli (including the cuing sound) were spatialized using three different levels of naturalness and richness: 1) generic ITDs only, 2) individualized IIDs, or 3) individualized HRTFs containing all of the naturally occurring spatial cues a listener experiences in the everyday world. We show that behavioral performance is better when listeners hear natural, individualized spatial cues than when they hear impoverished cues. Importantly, only natural spatial cues yield significant attentional modulation of P1 amplitudes. Moreover, induced alpha activity is less robust and poorly lateralized with isolated spatial cues compared to rich, natural spatial cues.

2 Materials and Methods

2.1 Subjects

Twenty-one paid volunteers and one author within the age of 18–42 years (M = 22.9, SD = 5.5; 12 females, 10 males) participated in this study. None of the subjects had audiometric thresholds greater than 20 dB for frequencies from 250 Hz to 8 kHz. All participants gave informed consent as approved by the Boston University Institutional Review Board. Two subjects were withdrawn from the study due to the inability to perform the task (percentage of correct response less than 30% after training), and two subjects were removed during EEG data preprocessing due to excessive artifacts. Therefore 18 subjects remained for further analysis (N = 18).

2.2 Stimuli and Procedure

The sound stimuli consisted of consonant-vowel syllables (/ba/, /da/, & /ga/), each 0.4 s in duration. These syllables were recorded from three talkers that naturally differed in fundamental frequency (F0). Details on stimulus are provided in Stimulus Presentation. Cue and stimuli were presented via earphones (ER-2, Etymotic Research, Inc.) and spatialized to approximately ±30° azimuth (0° elevation). Three different spatialization conditions were used: HRTF, IID, and ITD. In the HRTF condition, individualized HRTFs, providing natural combinations of ITDs, IIDs, and spectral cues, were used (See Individual HRTF Measurement for measurement methods). In the IID condition, ITDs were removed from the individualized HRTFs by computing minimum-phase representations of the filters via removing the non-causal part of the cepstrum. Hence, the IID and HRTF conditions provided the same monaural magnitude spectra and thus the same energetic advantage of the ear ipsilateral to the target. In the ITD condition, spatialization was based on simply delaying the signal presented to the contralateral ear by 300 µs, thus providing no energetic advantage to the ipsilateral ear.

The auditory cue was a single syllable /ba/ spoken by a low-pitch male voice (F0 = 91 Hz, estimated by Praat software) (Boersma, 2001). The following target and distractor streams both consisted of three syllables randomly chosen out of the set of three syllables (with replacement). The target stream was spoken by either a female (F0 = 189 Hz) or a high-pitch male talker (F0 = 125 Hz), and the distractor stream was spoken by the other talker different than the target stream. The first syllable of the target and distractor sound overlapped in time, while the latter two syllables were separated by 200 ms, onset to onset (Figure 26). To avoid engagement of temporal attention rather than spatial attention, the assignment of the target stream being leading or lagging was equally distributed across trials. In the leading stream, the onsets of all three syllables were separated by 400 ms; in the lagging stream, the onsets of the first and the second syllable were separated by 600 ms, whereas those of the second and the third syllable were separated by 400 ms. All sound stimuli were presented at a sound pressure level of approximately 75 dB.

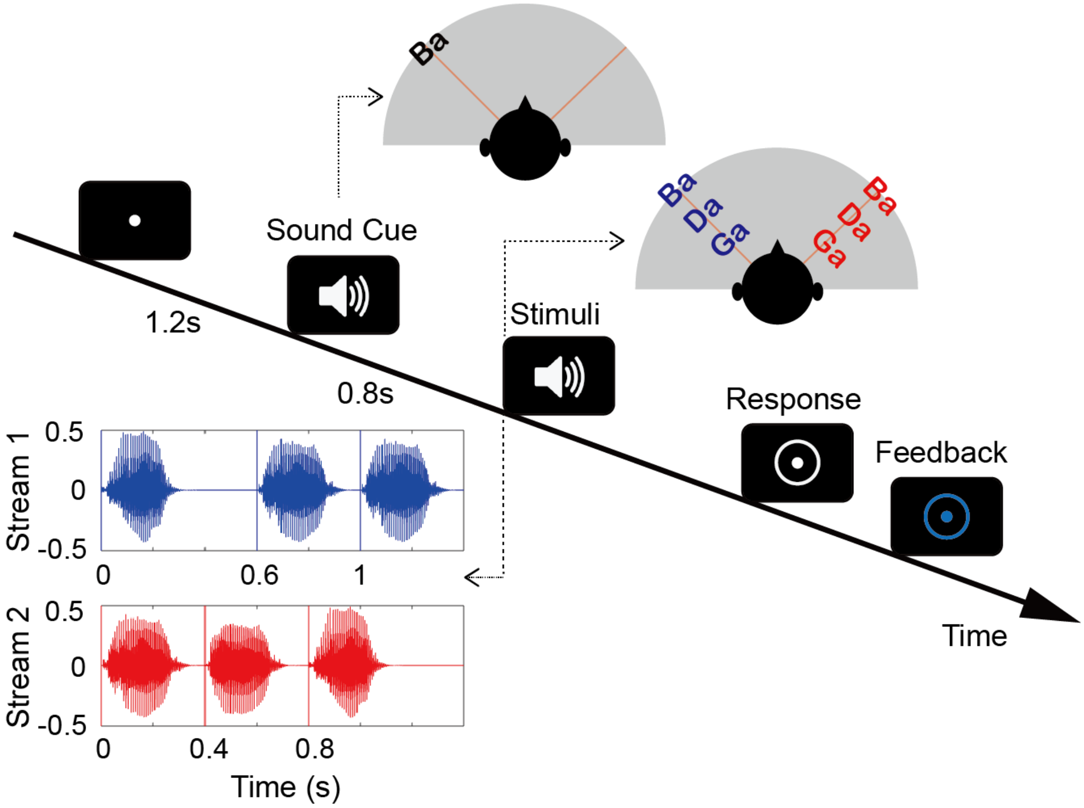

Figure 26. Auditory spatial attention task with two competing streams was used to assess the consequence of impoverished auditory spatial cues on neural proxies of attention control. An auditory cue was presented first from the location of the upcoming target stream, processed by the same spatialization scheme as the upcoming mixture. Following the cue, the competing streams began, one from around +30° the other from around -30° azimuth. Listeners were asked to recall the syllable sequence presented from the cued side. The first syllables of both streams were temporally aligned; however, the latter two syllables in the competing streams were staggered, enabling us to isolate neural responses to each. Feedback was provided after every trial.

2.3 Task

Subjects performed a spatial attention task in a Posner paradigm (Figure 26) (Posner et al., 1980). Sound spatialization was realized by one of the three spatialization conditions fixed within trials but pseudo-randomized across trials. Subjects were instructed to fixate their gaze on a dot at the center of the screen at the beginning of each trial. The fixation dot lasted for 1.2 s before an auditory cue was presented. The auditory cue came from either left or right indicating where the target sound would come from. A target sound started 0.8 s later from the cued location. At the same time a distractor sound started from the opposite location of the target sound. Subjects were asked to report the syllable sequence of the target sound by pressing a keyboard after the sounds finished and a response cue was shown. Feedback about whether or not they correctly reported the syllables was given at the end of every trial.

Each subject performed 450 randomized trials of this task, divided into 9 blocks each consisting of 50 trials. In total, every subject performed 150 trials for each of the three sound spatialization conditions (75 trials attending left and 75 trials attending right). Prior to the test sessions, all participants received a practice session to get familiarized with the task. Participants with a percentage of correct response lower than 30% after 3 blocks of training (50 trials per block) were excluded from the study.

2.4 EEG Acquisition and Preprocessing

32-channel scalp EEG data was recorded (Activetwo system with Activeview acquisition software, Biosemi B.V.) in a sound proof booth (Eckel Industries, Inc.) while subjects were performing the task. Two additional reference electrodes were placed on the earlobes. Horizontal eye movements were recorded by two electrooculography (EOG) electrodes placed on the outer canthi of each eye. Vertical eye movement was recorded by one EOG electrode placed below the right eye. The timing of stimulus was controlled by Matlab (MathWorks) with Psychtoolbox (extension 3) (Brainard, 1997; Delorme and Makeig, 2004b).

EEG preprocessing was conducted with EEGlab toolbox in Matlab. EEG data were corrected against the average of the two reference channels. Bad channels were marked by manual selection during recording and automatically detected based on joint probability measures of EEGlab (Arnaud Delorme and Swartz, 2004). EEG signals were then down-sampled to 256 Hz and epochs containing responses to individual trials were extracted. Each epoch was baseline corrected against 100 ms prior to the cue onset by removing the mean of the baseline period from the whole trial. ICA artifact rejection was performed with EEGlab to remove components of eye movements, blinks, and muscle artifacts. The maximum number of independent components rejected for each subject was five. After ICA rejection, bad channels were removed and interpolated. Trials with a maximum absolute value over 80 µV were rejected (Delorme et al., 2007). Two subjects with excessive artifacts were removed from further EEG analysis because less than 50% of trials remained after thresholding. For the rest of the 18 subjects, at least about two thirds of the trials (minimum was 48 out of 75 trials) remained for each condition after artifact rejection. Trial numbers were equalized within and across subjects by randomly selecting the minimum number of available trials (N = 48) for each condition across the whole recording session.

2.5 Data analysis

Behavioral performance was quantified by the percentage of correct responses for every one of the three syllables in the target stream and each spatialization condition. Behavioral results were collapsed across the attend-left and attend-right trials. The percentages of correct response were then normalized by logit transformation before parametric statistical testing was performed on the resulting data.

ERP responses were evaluated for the second syllable of the target sound and distractor sound, respectively. The reason we looked at the second syllable only is that 1) the first syllable of the target and distractor aligned in time and therefore the ERPs were inseparable, and 2) the ERP amplitude in response to the third syllable was small, and therefore more contaminated by noise. ERP components were then extracted from the time series data. The preprocessed data (details see EEG Preprocessing Procedures) were bandpass filtered from 0.5 to 20 Hz by a finite impulse response filter with Kaiser window design ($\beta = 7.2$, $n = 1178$). Data from four fronto-central channels (Cz, Fz, FC1, and FC2) were averaged to get the auditory ERP response. We picked these four channels a priori because auditory ERP responses in sensor space are largest in the fronto-central area of the scalp. To quantify the amplitudes of ERP components, the maximum value within the window of 50 to 100 ms after the second syllable onset was taken to be the P1 amplitude; the minimum value within the window of 100 to 180 ms after the second syllable onset was calculated to be the N1 amplitude. The values extracted from the selected windows were calculated for each channel and plotted onto a 2D scalp map to generate topography plots. The values of the ERP components from the four selected channels were then averaged and compared across different spatialization conditions.

To get the amplitude of alpha oscillation, the preprocessed EEG data was bandpass filtered to the alpha range (8 to 14 Hz) before a Hilbert transform was applied. The magnitude of the resulting data was taken as the extracted alpha power envelope. To get induced alpha power, the alpha power was calculated for single trials first and then averaged across trials (Snyder and Large, 2005). The time course of alpha power was baseline corrected against 700 ms before the auditory cue onset. GFP (Murray et al., 2008; Skrandies, 1990) constitutes the spatial standard deviation across all scalp electrodes; it has been used as a measurement to quantify the amount of alpha variation across the scalp (Lim et al., 2015). We calculated the time courses of alpha GFP by taking the standard deviation of all electrodes. To quantify the degree of alpha lateralization based on direction of attention, we calculated the Attentional Modulation Index (AMI) of alpha power, defined as the alpha power difference between attended left and attended right trials divided by the overall alpha power (Wöstmann et al., 2016a). The AMI of alpha was calculated for each time point, yielding the time course of AMI for each spatialization condition. We then averaged the alpha AMI of each spatialization condition over the 800 ms immediately before stimulus onset (-800 ms to 0 ms, re: onset). This is the period where subjects have been cued to orient their spatial attention in preparation for the target sound, but before the speech streams begin. Scalp topographies of the preparatory alpha AMI were plotted for each condition. Hemispheric lateralization of alpha AMI was further compared across spatialization conditions and evaluated as the difference between the left hemisphere and the right hemisphere.

For hypothesis testing, we compared group level differences between conditions with paired t-tests. Due to the relatively small number of subjects (N=18), the Lilliefors test was performed prior to statistical testing to check normality of the data. Data was considered normally distributed at P>0.05.

3 Results

3.1 Natural spatial cues facilitate behavioral performance

We compared the percentage of correctly recalling each syllable of the target stream across the three spatialization conditions (Figure 27). Paired t-tests were performed on the logit transformed percentages. For the first syllable, where the target sound and the distractor sound overlapped in time, subjects were least accurate in discriminating the target sound in the ITD condition compared to the IID condition ($t_{(17)} = 5.15$, $P < 0.001$) and HRTF condition ($t_{(17)} = 7.09$, $P < 0.001$). However, no statistically significant difference was observed between IID and HRTF conditions for the first syllable ($t_{(17)} = 1.4$, $P = 0.18$). For the second and the third syllable, where target and distractor streams occurred staggered in time, subjects performed significantly better in the HRTF condition than in both the ITD condition (2nd syllable: $t_{(17)} = 2.78$, $P = 0.013$; 3rd syllable: $t_{(17)} = 2.68$, $P = 0.016$) and the IID condition (2nd syllable: $t_{(17)} = 3.15$, $P < 0.01$; 3rd syllable: $t_{(17)} = 2.15$, $P = 0.046$). There was no significant difference between the ITD and IID conditions for the two staggered syllables (2nd syllable: $t_{(17)} = 0.49$, $P = 0.63$; 3rd syllable: $t_{(17)} = 1.74$, $P = 0.099$).

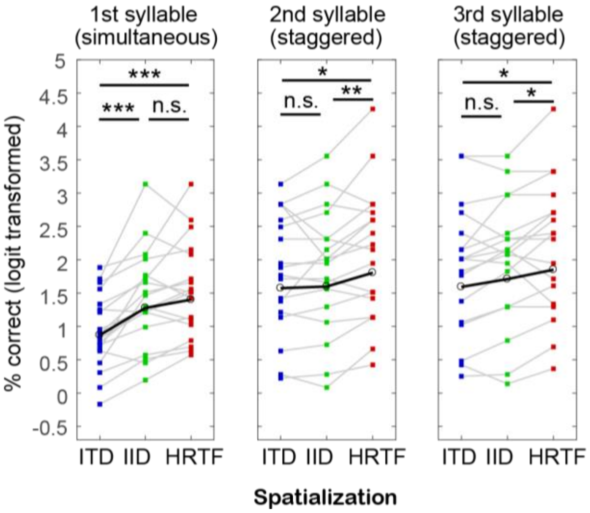

*Figure 27. Listeners' (N = 18) recall performance was evaluated for every syllable and different spatialization conditions. Sounds were spatialized either based on generic ITDs, individualized IIDs, or the natural combination of ITDs, IIDs, and spectral cues in individualized HRTFs. Behavioral advantages of having more consistent spatial information were significant but small. * P < .05; ** P < .001; *** P < .0001*

3.2 Impoverished spatial cues affect attentional modulation of ERPs

Figure 28A shows the ERPs evoked by the onset of the second syllable of the attended target sound and the unattended distractor sound, aligning the onsets of the target and distractor syllables to 0 s to allow direct comparison. Stimulus onsets elicited a fronto-central positivity (P1) between 50 to 100 ms followed by a negativity (N1) between 100 to 180 ms (Figure 28A-B). The amplitudes of these two components were extracted and the difference between attended stimuli (target sound) and unattended stimuli (distractor sound) was calculated in order to quantify attentional modulation for both the P1 and N1 components (Figure 28C).

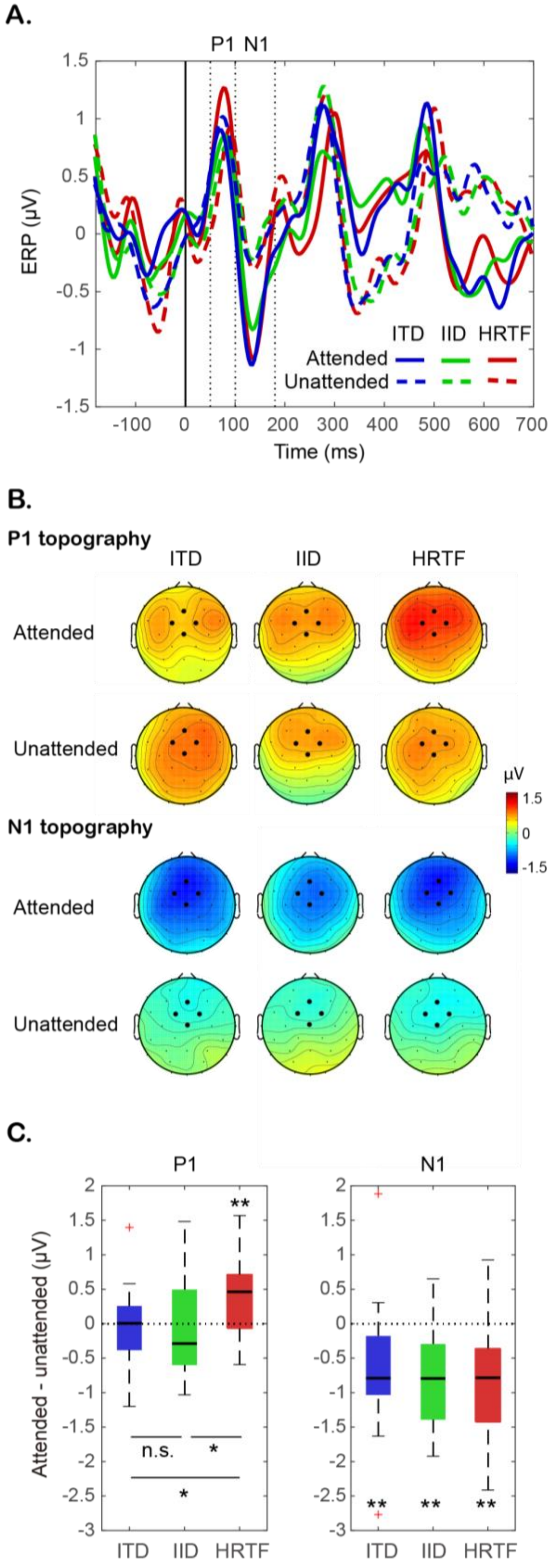
A.
P1 N1
1.5
1
0.5
0
-0.5
-1
-1.5
ERP (μV)
-100
0
100
200
300
400
500
600
700
Time (ms)
ITD IID HRTF
Attended
Unattended
B.
P1 topography
ITD
IID
HRTF
Attended
Unattended
μV
1.5
0
-1.5
N1 topography
Attended
Unattended
C.
P1
N1
Attended - unattended (μV)
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
-3
**
n.s.
*
*
ITD
IID
HRTF
** ** **
ITD
IID
HRTF

Figure 28. P1 amplitudes were only modulated by the direction of attention in the HRTF condition, whereas N1 amplitudes were modulated equivalently across spatialization conditions (N = 18). ***A.*** *ERP waveforms at fronto-central electrodes were compared between the attended target stream and the unattended distractor stream for every spatialization condition. The P1 time range was defined between 50 ms and 100 ms, and the N1 time range was defined between 100 ms and 180 ms.* ***B.*** *Most topographies of both ERP components show maxima at the fronto-central sites (black dots) used for evaluation.* ***C.*** *The modulation strength of ERP components was assessed by the amplitude differences between attended and unattended streams.*

We conducted repeated measures one-way ANOVA and found statistically significant effect of spatial cue on P1 amplitude modulation ($F_{(2,34)} = 3.34$, $P = 0.047$). P1 responses to attended stimuli were significantly larger than to unattended stimuli in the HRTF condition ($t_{(17)} = 3.12$, $P < 0.01$). However, no significant attentional modulation was found in the ITD ($t_{(17)} = 0.50$, $P = 0.62$) and IID conditions ($t_{(17)} = 0.06$, $P = 0.95$). Consistent with this, a direct comparison of the amount of modulation (difference in ERP magnitudes for attended versus unattended sounds) found that attentional modulation was significantly larger in the HRTF condition than in the ITD ($t_{(17)} = 2.48$, $P = 0.024$) and IID conditions ($t_{(17)} = 2.29$, $P = 0.035$); however, modulation did not differ significantly between the ITD and IID conditions ($t_{(17)} = 0.27$, $P = 0.79$) (Figure 28C).

In all three spatialization conditions, the N1 amplitude was modulated by spatial attention. An attended sound elicited larger N1 amplitudes than when a sound was ignored in all three conditions (ITD: $t_{(17)} = 3.01$, $P < 0.01$; IID: $t_{(17)} = 4.12$, $P < 0.001$; HRTF: $t_{(17)} = 3.56$, $P < 0.01$). The magnitude of this N1 modulation did not differ significantly across the three spatialization conditions (ITD vs IID: $t_{(17)} = 0.32$, $P = 0.76$; ITD vs HRTF: $t_{(17)} = 0.21$, $P = 0.84$; HRTF vs IID: $t_{(17)} = 0.13$, $P = 0.90$) (Figure 28C).

3.3 Alpha oscillation power shows less attentional modulation with impoverished spatial cues

To investigate the effect of spatialization on attentional control, we analyzed the power in alpha oscillations during the attentional preparation period (-800 ms to 0 ms), a time period in which listeners know where to orient spatial attention based on the preceding acoustic cue, but before the sound mixture of competing streams begins. We averaged the power in alpha across all trials for each spatialization condition, regardless of where spatial attention was focused, to get a measure of the total engagement of alpha activity. We then compared relative power for different attentional directions. Averaged across direction of attentional focus, we calculated the time courses of alpha global field power (GFP, Figure 29A) and compared within-subject differences of the temporal average within the preparatory time period across spatialization conditions (Figure 29B). We conducted repeated measures one-way ANOVA and found statistically significant effect of spatial cue on GFP amplitude ($F_{(2,34)} = 5.26$, $P = 0.01$). We found significantly larger alpha GFP differences in the HRTF condition than in the other two conditions (HRTF vs ITD: $t_{(17)} = 2.75$, $P = 0.014$; HRTF vs IID: $t_{(17)} = 2.63$, $P = 0.017$). No significant difference was found between the ITD and IID conditions ($t_{(17)} = 0.021$, $P = 0.98$).

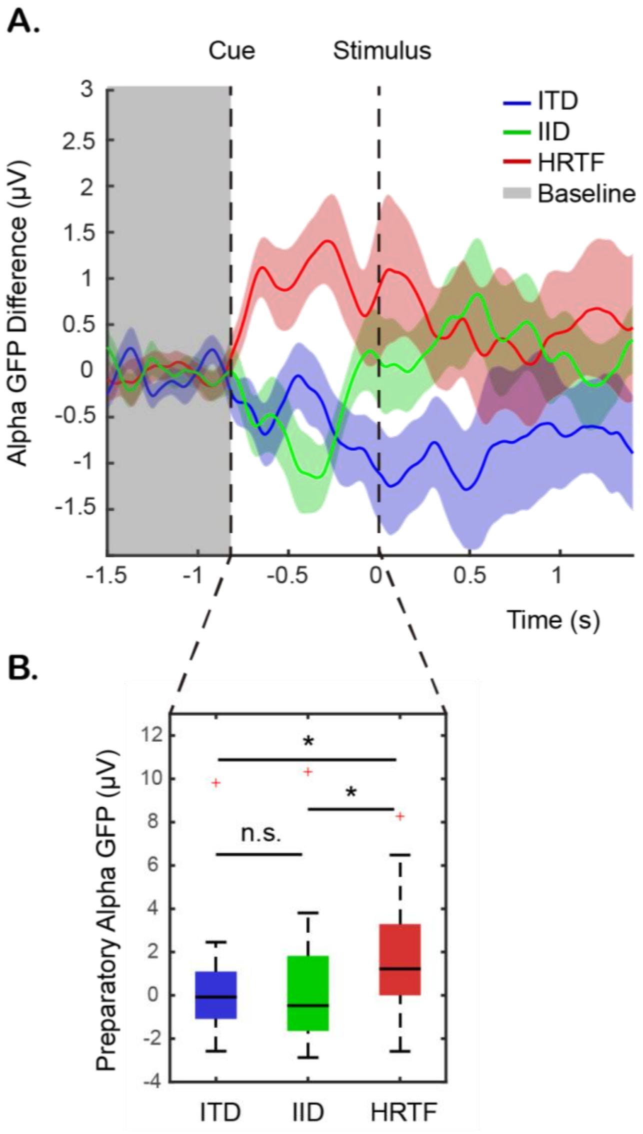

Figure 29. Within-subject differences in alpha-band GFP are larger in the HRTF condition, especially during the preparatory time window (after the sound cue but before the first syllables of the competing streams). A. Waveforms of the average (± SEM) GFP differences are shown during the baseline period, preparatory phase, and stimulus phase with stream competition. B. The temporal average of the preparatory alpha GFP difference is larger for the HRTF condition.

We next explored whether the lateralization of alpha with the spatial focus of attention was greater with more natural spatial cues than with impoverished cues. To this end, for each spatialization condition we calculated the alpha power “attentional modulation index” (AMI), defined as the alpha power difference between attend left and attend right trials divided by the overall alpha power (Wöstmann et al., 2016a). Calculated in this way, the AMI is expected to be positive in left and negative in right parietal channels. Scalp topographies of the resulting AMI are shown in Figure 30. The alpha AMI was then averaged within the left and the right parietal channels (see Methods). Statistical comparisons between left- and right-hemisphere AMIs were made within conditions to quantify lateralization of alpha power. Statistically significant hemispheric differences were found in the HRTF condition only ($t_{(17)} = 3.09$, $P < 0.01$), not in the ITD ($t_{(17)} = 1.29$, $P = 0.21$) or IID conditions ($t_{(17)} = 0.15$, $P = 0.88$). To compare alpha lateralization directly across conditions, the hemispheric difference in alpha AMI was calculated by subtracting average right-hemisphere AMI from the average left-hemisphere AMI. Direct comparison of conditions shows that alpha lateralization was significantly greater in the HRTF condition than in the ITD condition ($t_{(17)} = 2.18$, $P = 0.043$); alpha lateralization in the IID condition did not differ significantly from that of either the HRTF condition ($t_{(17)} = 1.82$, $P = 0.086$) or the ITD condition ($t_{(17)} = 0.72$, $P = 0.48$).

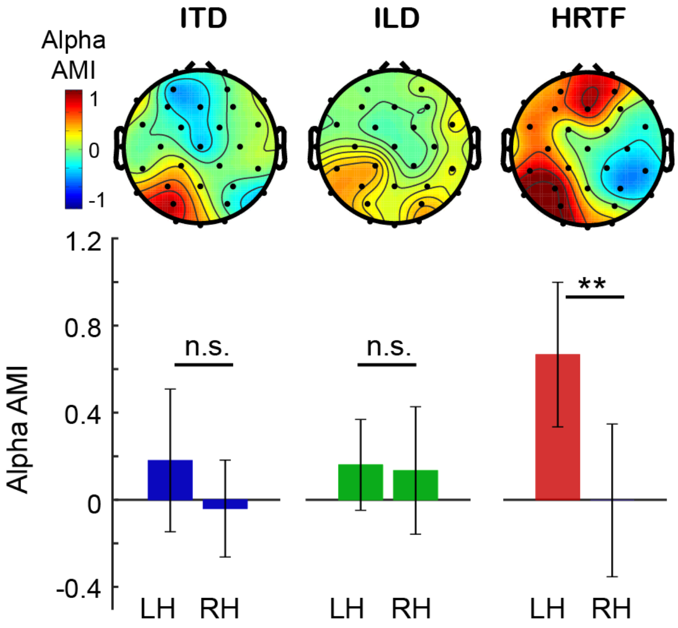

*Figure 30. Attentional modulation of alpha activity was lateralized to the hemisphere ipsilateral to the target stream only in the HRTF condition. AMI topographies and hemispheric averages are shown for every spatialization condition (N = 18). ** P < .001*

Impoverished spatial cues lead to worse behavioral performance, smaller P1 modulation, reduced modulation of preparatory alpha power GFP, and reduced lateralization of alpha power with attentional focus, confirming our hypothesis that impoverished spatial cues impaired engagement of spatial attention.

4 Discussion

Behaviorally, we found that impoverished spatial cues impair performance on an auditory spatial attention task in a multi-talker scene. We used objective electrophysiological measures to assess whether the naturalness and richness of spatial cues also impacts how strongly auditory spatial attention modulates brain responses. We found that impoverished spatial cues reduce the strength of the evoked and induced neural signatures of attentional control. Specifically, evoked P1 amplitudes and induced alpha oscillatory power showed less attentional modulation for sound stimuli with impoverished spatial cues compared to when spatial cues were tailored to recreate the natural, rich experience of individual listeners.

4.1 Impoverished spatial cues result in less neural modulation during selective attention.

We investigated attentional modulation of four established neural signatures of selective attention: evoked P1 and N1 amplitudes and induced power and lateralization of alpha oscillation. While attentional modulation of N1 amplitude was observed in all conditions, attentional modulation of the earlier P1 amplitude was not observed or was significantly weaker in the impoverished cue conditions compared to the natural cue condition. Similarly, we found less preparatory alpha power activity in the impoverished spatial cue conditions than in the natural cue condition, reflected by two indexes quantifying the amount of spatial variability of alpha power: alpha GFP (Figure 29) and AMI (Figure 30). In the ITD and IID conditions, although there was a hint of preparatory alpha lateralization over parietal sensors, the amount of lateralization was significantly smaller than in the HRTF condition and did not reach statistical significance. Preparatory alpha activity during spatial attention tasks has been well documented to form a specific lateralization pattern in both vision and audition (S Banerjee et al., 2011; Kelly, 2006c; Paul Sauseng et al., 2005; Worden et al., 2018), which is thought to be evidence of a preparatory information-gating mechanism (John J. Foxe and Snyder, 2011; Jensen and Mazaheri, 2010a; Klimesch et al., 2007b; Klimesch, 2012a). In vision, alpha lateralization has been observed to increase with the laterality of attention focus (Rihs et al., 2007a; Samaha et al., 2015), reflecting an inhibition pattern topographically specific to attention focus. Moreover, evidence for active top-down control of the phase of alpha oscillation during visual spatial attention suggests that alpha oscillatory activity represents active engagement and disengagement of the attentional network (Samaha et al., 2016b). In addition, a previous somatosensory study has revealed that the alpha lateralization is positively correlated to pre-stimulus cue reliability, further suggesting that alpha lateralization reflects the top-down control in order to optimize the processing of upcoming stimuli (Saskia Haegens et al., 2011). Although relatively few studies have investigated alpha activity in audition, studies suggest that alpha control mechanisms are supra-modal rather than sensory specific (S Banerjee et al., 2011). In the current experiment, a pre-stimulus auditory cue directed listeners where to focus attention in an upcoming sound mixture. The cue was spatialized using the same auditory features used to spatialize the stream mixture. Our results thus suggest that compared to stimuli with natural spatial cues, stimuli featuring only ITDs or only IIDs are less reliable in directing attentional focus, producing weaker engagement of spatial attention and reduced attentional modulation of neural responses.

Consistent with the idea that impoverished spatial cues lead to weaker engagement of spatial attention, we found that the P1 ERP component was modulated by attention only with natural spatial cues, not with impoverished cues; this result is consistent with a weak spatial representation failing to engage attentional modulation of early sensory responses (Figure 28). Our finding that attentional focus leads to a modulation of P1 amplitude for natural spatial cues is consistent with previous reports of effects of attention on the P1 amplitude observed in previous spatial attention studies across sensory modalities [auditory: (Giuliano et al., 2014); visual: (Hillyard and Anllo-Vento, 1998; Hopfinger et al., 2004)]. Past studies agree that P1 modulation reflects an early sensory inhibition mechanism related to suppression of task-irrelevant stimuli. Although debates remain as to whether P1 modulation results from bottom-up sensory gain control (Hillyard and Anllo-Vento, 1998; Luck, 1995; H A Slagter et al., 2016) or for some top-down inhibitory process (Freunberger et al., 2008; Klimesch, 2011), it is generally accepted in visual spatial studies that greater P1 amplitude modulation is associated with greater inhibition of to-be-ignored stimuli (J.W. Couperus and Mangun, 2010; Hillyard and Anllo-Vento, 1998; Klimesch, 2012a). Interestingly, attentional modulation of auditory P1 has been found to be positively correlated to visual working memory capacity, a result that was used to suggest that stronger P1 modulation reflects better attentional control of the flow of sensory information into working memory (Fukuda and Vogel, 2009; Giuliano et al., 2014). Our result is consistent with the hypothesis that P1 modulation directly reflects attentional control. Specifically, impoverished spatial cues likely produce a “muddy” representation of auditory space that supports only imprecise, poorly focused top-down spatial attention. The resulting lack of control and specificity of spatial auditory attention results in early P1 responses that are unmodulated by attentional focus.

N1 modulation is well documented as a neural index of attentional control (Choi et al., 2013; Hillyard et al., 1998; Stevens et al., 2008; Wyart et al., 2012). The attentional modulation of N1 is thought to reflect attentional facilitation rather than inhibition (J. W. Couperus and Mangun, 2010; Marzecová et al., 2018; H. A. Slagter et al., 2016). In contrast to preparatory alpha and P1, we found that the later N1 evoked response was modulated similarly, regardless of the richness and naturalness of spatial cues.

Due to the robustness and relatively large amount of modulation, changes in auditory N1 amplitude have been used as a biomarker and a primary feature for classification of attentional focus (Blankertz et al., 2011; Schreuder et al., 2011); see also recent work on decoding attentional focus for running speech using the correlation between neural responses and the power envelope of the speech streams: (Chait et al., 2010; Mesgarani and Chang, 2012; Rimmele et al., 2015)). However, there is little known about how N1 amplitudes reflects the processing of different spatial cues during auditory spatial attention. Previous studies have revealed different N1 topographies during ITD and IID processing, leading to the conclusion that ITD and IID are processed by different neural populations in the auditory cortex (Johnson and Hautus, 2010; Tardif et al., 2006; Ungan et al., 2001). However, debates remain about whether this difference in topography depends on perceived laterality, instead of different neural populations specialized for processing different spatial cues. Results from a more recent study show that auditory N1 modulation does not differ across spatial cue conditions, indicating integrated processing of sound locations in auditory cortex regardless of cues (Salminen et al., 2015). In the current study, N1 modulation did not differ across the three spatialization conditions. Thus, our results support the idea that the same cortical neural population is responsible for processing different binaural spatial cues.

4.2 Behavioral disadvantages associated with impoverished spatial cues are modest and depend on sound stimulus characteristics

Despite the influence of spatial cue richness on neural metrics, our behavioral results showed only small (albeit significant) behavioral differences between impoverished spatial cues and natural, individualized spatial cues (Figure 27). In line with previous studies that observed greater spatial release from masking with combined spatial cues compared to with isolated cues (Culling et al., 2004; Ellinger et al., 2017), accuracy was best in the HRTF condition. The small accuracy improvement over using impoverished cues is seen consistently across subjects. In the first syllable where the target and distractor streams overlap in time, the HRTF condition yielded a 13% increase in accuracy over the ITD condition, but is comparable to performance in the IID condition. In the two staggered syllables, accuracy in the HRTF condition is greater than in the ITD and IID conditions by only about 6% and 1%, respectively. These differences in behavioral performance across syllables suggest that the characteristics of sound stimuli influence the difficulty of the task and may influence the behavioral advantages of having richer, more robust spatial cues (Kidd et al., 2010). Concordantly, a previous study with complex tone stimuli has shown much larger differences in behavioral performance of up to 20% (Schröger, 1996), whereas studies presenting speech stimuli in a multi-talker environment found no behavioral advantage of having combined cues compared to impoverished cues (Glyde et al., 2013). These behavioral discrepancies, in combination with our neural findings, indicate that behavioral performance alone is not a sensitive metric for determining whether cortical networks controlling spatial selective attention are fully engaged.

Conclusion

Our results indicate that although impoverished spatial cues can support spatial segregation of speech in a multi-talker environment, they do not fully engage the brain networks controlling spatial attention and lead to weak attentional control. Previous auditory studies have provided evidence that impoverished spatial cues do not evoke the same neural processing mechanisms as natural cue combinations during localization tasks with single sounds (Callan et al., 2013; Getzmann and Lewald, 2010; Leino et al., 2007; Palomäki et al., 2005). The current study extends these findings, demonstrating that the efficacy of higher-level cognitive processing, such as deployment of auditory selective attention, also depends on the naturalness of spatial cues. Poor attentional control was reflected in limited modulation of neural biomarkers of attentional processes. These findings suggest that the many past auditory attention studies using impoverished spatial cues may have underestimated the robust changes in cortical activity associated with deployment of spatial auditory attention in natural settings. Although impoverished auditory spatial cues can allow listeners to deploy spatial attention effectively enough to perform well in simple acoustic scenes, noisy, complex listening environments like those encountered in everyday environments pose greater challenges to attentional processing. In natural settings, spatial attention may fail unless attentional control networks are fully engaged. Thus, these results demonstrate the importance of preserving rich, natural spatial cues in hearing aids and other assistive listening devices.

SUMMARY AND DISCUSSION

Results from this thesis study furthered the understanding of cortical activities underlying auditory attention based on spatial and non-spatial physical characteristics. Specifically, it was demonstrated in both correlative and causal aspects that alpha oscillation plays differential functional roles during different forms of auditory attention.

In Chapter 1, we aimed to study the correlative alpha activities during auditory spatial attention. We demonstrated that similar to visual spatial attention, auditory spatial attention modulates parietal alpha oscillation so its lateralization pattern is topographically specific to spatial attention focus. This modulation is limited to alpha in parietal cortex. As shown in alpha peak frequency analysis, multiple alpha generators exists during this attentional task, and alpha in the central cortex does not show topographic specificity to spatial attention focus. Consistent with visual spatial attention study results (Rihs et al., 2007b; Samaha et al., 2016c), we provided further evidence that auditory and visual spatial attention shares parietal spatial control network (Green et al., 2011).

In Chapter 2, we further investigated in whether parietal alpha activities modulation is specific to auditory spatial attention. With a task design that allows direct comparison between auditory spatial and non-spatial attention, we investigated into dividing attentional mechanism underlying these two forms of attention. Our results demonstrated that spatial and non-spatial auditory attention differs in behavioral strategy and neural mechanisms employed. Alpha oscillation in the parietal cortex is modulated during divided spatial attention but not non-spatial attention, suggesting that auditory “where” and “what” pathways involve different cortical networks (Larson and Lee, 2014).

Besides correlative relationships between alpha and auditory attention, in Chapter 3 we showed a causal role alpha plays during auditory attention. With HD-tACS, we stimulated the right hemisphere of listeners while they were performing a spatial and non-spatial attention task. We found that increased alpha on the right hemisphere disturbs attention focused on the contralateral left hemifield. This stimulation effect is not only hemispheric specific, but also frequency specific (6H theta oscillation does not show any stimulation effects), task specific (tasks does not require continuous alpha lateralization does not show stimulation effect) and specific to forms of auditory attention (non-spatial attention performance not affected). These results provide evidence of a causal function of alpha during auditory spatial attention.

With the existing knowledge of alpha as a neural signature of auditory spatial attention, in Chapter 4 we applied this knowledge to use alpha as a quantitative measurement of the level of engagement of auditory spatial attention network. We studied auditory spatial attention under three spatial cue conditions: ITD only, ILD only and natural individualized HRTF. We found that impoverished spatial cues do not engage the spatial attention network as much as the natural spatial cues. Our results have the important implication that auditory attention studies using impoverished cues to deliver spatialized sound effect is likely to underestimate the amount of neural response.

Together, our results provided evidence in both the correlative and causal aspect that alpha oscillation is a functional mechanism during auditory spatial attention. Depending on its generator, the role of alpha oscillation differs when auditory attention is focus on spatial or non-spatial physical attributes of the sound. With alpha being an important biomarker of auditory spatial attention, it poses promising potential to be employed to develop brain computer interfaces and clinical diagnosis. Further studies on the generators of alpha during auditory non-spatial attention and the neurophysiological underpinning of alpha would be of great interest to further understand this neural phenomenon.

BIBLIOGRAPHY

Ahveninen, J., Hämäläinen, M., Jääskeläinen, I.P., Ahlfors, S.P., Huang, S., Lin, F.-H., Raij, T., Sams, M., Vasios, C.E., Belliveau, J.W., 2011. Attention-driven auditory cortex short-term plasticity helps segregate relevant sounds from noise. Proceedings of the National Academy of Sciences of the United States of America. 108, 4182–4187. https://doi.org/10.1073/pnas.1016134108

Alho, K., Medvedev, S. V, Pakhomov, S. V, Roudas, M.S., Tervaniemi, M., Reinikainen, K., Zeffiro, T., Naatanen, R. 1999. Selective tuning of the left and right auditory cortices during spatially directed attention. Brain Research. Cognitive Brain Research. 7(3), 335–341.

Alho, K., Töttölä, K., Reinikainen, K., Sams, M., Näätänen, R., 1987. Brain mechanism of selective listening reflected by event-related potentials. Electroencephalography and Clinical Neurophysiology. Evoked Potentials 68, 458–470. https://doi.org/10.1016/0168-5597(87)90057-8

Altman, J.A., Vaitulevich, S.P., Shestopalova, L.B., Varfolomeev, A.L., 2005. Mismatch negativity evoked by stationary and moving auditory images of different azimuthal positions. Neuroscience Letters. 384, 330–335. https://doi.org/10.1016/j.neulet.2005.05.002

Antal, A., Paulus, W., 2013. Transcranial alternating current stimulation (tACS). Frontiers in Human Neuroscience. 7, 317. https://doi.org/10.3389/fnhum.2013.00317

Arnaud Delorme, S.M., Swartz, 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of Neuroscience Methods. 134(1) 9–21.

Arnott, S.R., Binns, M.A., Grady, C.L., Alain, C., 2004. Assessing the auditory dual-pathway model in humans. NeuroImage 22, 401–408. https://doi.org/10.1016/j.neuroimage.2004.01.014

Atilgan, I., 1998. Basin??li su reakt??rlerinde yakit ??ubuklarinda tek boyutlu isi iletiminin analizi. J. Fac. Eng. Archit. Gazi Univ. 13, 141–152.

Banerjee, S., Snyder, A.C., Molholm, S., Foxe, J.J., 2011. Oscillatory Alpha-Band Mechanisms and the Deployment of Spatial Attention to Anticipated Auditory and Visual Target Locations: Supramodal or Sensory-Specific Control Mechanisms? Journal of Neuroscience. 31, 9923–9932. https://doi.org/10.1523/JNEUROSCI.4660-10.2011

Banerjee, S., Snyder, A.C., Molholm, S., Foxe, J.J., 2011. Oscillatory Alpha-Band Mechanisms and the Deployment of Spatial Attention to Anticipated Auditory and Visual Target Locations: Supramodal or Sensory-Specific Control Mechanisms? Journal of Neuroscience. 31, 9923–9932. https://doi.org/10.1523/JNEUROSCI.4660-10.2011

Barnhardt, J., Ritter, W., Gomes, H., 2008. Perceptual load affects spatial and nonspatial visual selection processes: An event-related brain potential study. Neuropsychologia 46, 2071–2078. https://doi.org/10.1016/j.neuropsychologia.2008.02.007

Basar, E., 2012. A review of alpha activity in integrative brain function: Fundamental physiology, sensory coding, cognition and pathology. International Journal of Psychophysiology. 86, 1–24.

Baumgartner, R., Reed, D.K., Tóth, B., Best, V., Majdak, P., Colburn, H.S., Shinn-Cunningham, B., 2017. Asymmetries in behavioral and neural responses to spectral cues demonstrate the generality of auditory looming bias. Proceedings of the National Academy of Sciences of the United States of America. 114, 9743–9748. https://doi.org/10.1073/pnas.1703247114

Bendor, D., Wang, X., 2005. The neuronal representation of pitch in primate auditory cortex. Nature 436, 1161–1165. https://doi.org/10.1038/nature03867

Best, V., Ozmeral, E.J., Kopco, N., Shinn-Cunningham, B.G., 2008. Object continuity enhances selective auditory attention. Proceedings of the National Academy of Sciences of the United States of America. 105, 13174–13178. https://doi.org/10.1073/pnas.0803718105

Bestmann, S., de Berker, A.O., Bonaiuto, J., 2015. Understanding the behavioural consequences of noninvasive brain stimulation. Trends in Cognitive Sciences. 19, 13–20. https://doi.org/10.1016/j.tics.2014.10.003

Bidet-Caulet, A., Bertrand, O., 2005. Dynamics of a Temporo-Fronto-Parietal Network during Sustained Spatial or Spectral Auditory Processing. Journal of Cognitive Neuroscience. 17, 1691–1703. https://doi.org/10.1162/089892905774589244

Bikson, M., name, A., Rahman, A., 2013. Origins of specificity during tDCS: anatomical, activity-selective, and input-bias mechanisms. Frontiers in Human Neuroscience. 7, 688. https://doi.org/10.3389/fnhum.2013.00688

Blankertz, B., Lemm, S., Treder, M., Haufe, S., Müller, K.-R., 2011. Single-trial analysis and classification of ERP components — A tutorial. NeuroImage 56, 814–825. https://doi.org/10.1016/J.NEUROIMAGE.2010.06.048

Blauert, J., 1997. Spatial hearing: The psychophysics of human sound localization. Cambridge, MA: MIT Press.

Bodenmann, S., Rusterholz, T., Jaggi-Schwarz, K., Durr, R., Bachmann, V., Landolt, H.-P., Stoll, C., Geissler, E., 2009. The Functional Val158Met Polymorphism of COMT Predicts Interindividual Differences in Brain Oscillations in Young Men. Journal of Neuroscience. 29, 10855–10862. https://doi.org/10.1523/jneurosci.1427-09.2009

Boersma, P., 2001. Praat, a system for doing phonetics by computer. Glot International. 5, 341–345.

Brainard, D.H., 1997. The psychophysics toolbox. Spatial Vision. 10, 433–436.

Bressler, S., Masud, S., Bharadwaj, H., Shinn-Cunningham, B., 2014. Bottom-up influences of voice continuity in focusing selective auditory attention. Psychological Research. 78, 349–360. https://doi.org/10.1007/s00426-014-0555-7

Brignani, D., Ruzzoli, M., Mauri, P., Miniussi, C., 2013. Is Transcranial Alternating Current Stimulation Effective in Modulating Brain Oscillations? PLoS One 8, e56589. https://doi.org/10.1371/journal.pone.0056589

Brunetti, M., Della Penna, S., Ferretti, A., Del Gratta, C., Cianflone, F., Belardinelli, P., Caulo, M., Pizzella, V., Olivetti Belardinelli, M., Romani, G.L., 2008. A Frontoparietal Network for Spatial Attention Reorienting in the Auditory Domain: A Human fMRI/MEG Study of Functional and Temporal Dynamics. Cerebral Cortex 18, 1139–1147. https://doi.org/10.1093/cercor/bhm145

Buschman, T.J., Miller, E.K., 2007. Top-Down Versus Bottom-Up Control of Attention in the Prefrontal and Posterior Parietal Cortices. Science. 315(5820), 1860–1863.

Cai, Q., Van der Haegen, L., Brysbaert, M., 2013. Complementary hemispheric specialization for language production and visuospatial attention. Proceedings of the National Academy of Sciences of the United States of America. 110, E322–E330. https://doi.org/10.1073/pnas.1212956110

Callan, A., Callan, D.E., Ando, H., 2013. Neural correlates of sound externalization. NeuroImage 66, 22–27. https://doi.org/10.1016/j.neuroimage.2012.10.057

Capotosto, P., Babiloni, C., Romani, G.L., Corbetta, M., 2009. Transcranial magnetic stimulation of the human frontal eye field. Journal of Neuroscience. 29, 5863–72. https://doi.org/10.1523/JNEUROSCI.0539-09.2009

Chait, M., de Cheveigné, A., Poeppel, D., Simon, J.Z., 2010. Neural dynamics of attending and ignoring in human auditory cortex. Neuropsychologia 48, 3262–3271. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2010.07.007

Cherry, E.C., 1953. Some Experiments on the Recognition of Speech, with One and with Two Ears. Journal of the Acoustical Society of America. 25, 975–979. https://doi.org/10.1121/1.1907229

Choi, I., Rajaram, S., Varghese, L.A., Shinn-Cunningham, B.G., 2013. Quantifying attentional modulation of auditory-evoked cortical responses from single-trial electroencephalography. Frontiers in Human Neuroscience. 7, 115. https://doi.org/10.3389/fnhum.2013.00115

Cochin, S., Barthelemy, C., Lejeune, B., Roux, S., Martineau, J., 1998. Perception of motion and qEEG activity in human adults. Electroencephalography and Clinical Neurophysiology. 107, 287–295. https://doi.org/10.1016/S0013-4694(98)00071-6

Corbetta, M., Miezin, F.M., Dobmeyer, S., Shulman, G.L., Petersen, S.E., 1991. Selective and divided attention during visual discriminations of shape, color, and speed: functional anatomy by positron emission tomography. Journal of Neuroscience. 11, 2383–402.

Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain. Nature Reviews. Neuroscience. 3, 201–215. https://doi.org/10.1038/nrn755

Couperus, J.W., 2009. Implicit learning modulates selective attention at sensory levels of perceptual processing. Attention, Perception, & Psychophysics. 71, 342–351. https://doi.org/10.3758/APP.71.2.342

Couperus, J.W., Mangun, G.R., 2010. Signal enhancement and suppression during visual–spatial selective attention. Brain Research. 1359, 155–177. https://doi.org/10.1016/J.BRAINRES.2010.08.076

Culling, J.F., Hawley, M.L., Litovsky, R.Y., 2004. The role of head-induced interaural time and level differences in the speech reception threshold for multiple interfering sound sources. Journal of the Acoustical Society of America. 116, 1057–1065. https://doi.org/10.1121/1.1772396

Culling, J.F., Stone, M.A., 2017. Energetic Masking and Masking Release. pp. 41–73. https://doi.org/10.1007/978-3-319-51662-2_3

Cusack, R., Carlyon, R.P., Robertson, I.H., 2001. Auditory Midline and Spatial Discrimination in Patients with Unilateral Neglect. Cortex 37, 706–709. https://doi.org/10.1016/S0010-9452(08)70620-8

Cusack, R., Decks, J., Aikman, G., Carlyon, R.P. 2004, U., 2004. Effects of Location, Frequency Region, and Time Course of Selective Attention on Auditory Scene Analysis. Journal of Experimental Psychology: Human Perception and Performance.

30(4), 643–656 https://doi.org/10.1037/0096-1523.30.4.643

Dahmen, J.C., Keating, P., Nodal, F.R., Schulz, A.L., King, A.J., 2010. Adaptation to Stimulus Statistics in the Perception and Neural Representation of Auditory Space. Neuron 66, 937–948. https://doi.org/10.1016/J.NEURON.2010.05.018

Dai, L., Best, V., Shinn-Cunningham, B.G., 2018. Sensorineural hearing loss degrades behavioral and physiological measures of human spatial selective auditory attention. Proceedings of the National Academy of Sciences of the United States of America. 201721226. https://doi.org/10.1073/pnas.1721226115

Darwin, C.J., Brungart, D.S., Simpson, B.D., 2003. Effects of fundamental frequency and vocal-tract length changes on attention to one of two simultaneous talkers. Journal of the Acoustical Society of America. 114, 2913. https://doi.org/10.1121/1.1616924

Datta, A., Bansal, V., Diaz, J., Patel, J., Reato, D., Bikson, M., 2009. Gyri-precise head model of transcranial direct current stimulation: Improved spatial focality using a ring electrode versus conventional rectangular pad. Brain Stimulation. 2, 201–7, 207.e1. https://doi.org/10.1016/j.brs.2009.03.005

Dayan, E., Censor, N., Buch, E.R., Sandrini, M., Cohen, L.G., 2013. Noninvasive brain stimulation: from physiology to network dynamics and back. Nature Neuroscience. 16, 838–844. https://doi.org/10.1038/nn.3422

De Santis, L., Clarke, S., Murray, M.M., 2007. Automatic and intrinsic auditory “what” and “where” processing in humans revealed by electrical neuroimaging. Cerebral Cortex 17, 9–17. https://doi.org/10.1093/cercor/bhj119

De Schotten, M.T., Dell’Acqua, F., Forkel, S.J., Simmons, A., Vergani, F., Murphy, D.G.M., Catani, M., 2011. A lateralized brain network for visuospatial attention. Nature Neuroscience. 14, 1245–1246. https://doi.org/10.1038/nn.2905

Deans, J.K., Powell, A.D., Jefferys, J.G.R., 2007. Sensitivity of coherent oscillations in rat hippocampus to AC electric fields. Journal of Physiology. 583, 555–565. https://doi.org/10.1113/jphysiol.2007.137711

Degerman, A., Rinne, T., Salmi, J., Salonen, O., Alho, K., 2006. Selective attention to sound location or pitch studied with fMRI. Brain Research. 1077, 123–134. https://doi.org/10.1016/J.BRAINRES.2006.01.025

Degerman, A., Rinne, T., Srkk, A.-K., Salmi, J., Alho, K., 2008. Selective attention to sound location or pitch studied with event-related brain potentials and magnetic fields. European Journal of Neuroscience. 27, 3329–3341. https://doi.org/10.1111/j.1460-9568.2008.06286.x

Delorme, A., Makeig, S., 2004a. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of Neuroscience Methods 134, 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Delorme, A., Sejnowski, T., Makeig, S., 2007. Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. NeuroImage 34, 1443–1449. https://doi.org/10.1016/j.neuroimage.2006.11.004

Dosenbach, N.U., Fair, D.A., Miezin, F.M., Cohen, A.L., Wenger, K.K., T Dosenbach, R.A., Fox, M.D., Snyder, A.Z., Vincent, J.L., Raichle, M.E., Schlaggar, B.L., Petersen, S.E., 2007. Distinct brain networks for adaptive and stable task control in humans. Proceedings of the National Academy of Sciences of the United States of America. 104(26), 11073–11078

Ellinger, R.L., Jakien, K.M., Gallun, F.J., 2017. The role of interaural differences on speech intelligibility in complex multi-talker environmentsa). Journal of the Acoustical Society of America. 141, EL170-EL176. https://doi.org/10.1121/1.4976113

Farah, M.J., Wong, A.B., Monheit, M.A., Morrow, L.A., 1989. Parietal lobe mechanisms of spatial attention: Modality-specific or supramodal? Neuropsychologia 27, 461–470. https://doi.org/10.1016/0028-3932(89)90051-1

Foxe, J.J., Simpson, G. V., Ahlfors, S.P., 1998. Parieto-occipital ~1 0Hz activity reflects anticipatory state of visual attention mechanisms. Neuroreport. 9(17), 3929–3933.

Foxe, J.J., Snyder, A.C., 2011. The Role of Alpha-Band Brain Oscillations as a Sensory Suppression Mechanism during Selective Attention. Frontiers in Psychology. 2, 154. https://doi.org/10.3389/fpsyg.2011.00154

Freunberger, R., Höller, Y., Griesmayr, B., Gruber, W., Sauseng, P., Klimesch, W., 2008. Functional similarities between the P1 component and alpha oscillations. European Journal of Neuroscience. 27, 2330–2340. https://doi.org/10.1111/j.1460-9568.2008.06190.x

Frey, J.N., Mainy, N., Lachaux, J.-P., Muller, N., Bertrand, O., Weisz, N., 2014. Selective Modulation of Auditory Cortical Alpha Activity in an Audiovisual Spatial Attention Task. Journal of Neuroscience. 34, 6634–6639. https://doi.org/10.1523/JNEUROSCI.4813-13.2014

Freyman, R.L., Balakrishnan, U., Helfer, K.S., 2001. Spatial release from informational masking in speech recognition. Journal of the Acoustical Society of America. 109, 2112–2122. https://doi.org/10.1121/1.1354984

Fukuda, K., Vogel, E.K., 2009. Human Variation in Overriding Attentional Capture. Journal of Neuroscience. 29, 8726–8733. https://doi.org/10.1523/jneurosci.2145-09.2009

Gazzaley, A., Cooney, J.W., McEvoy, K., Knight, R.T., D'Esposito, M., 2005. Top-down enhancement and suppression of the magnitude and speed of neural activity. Journal of Cognitive Neuroscience. 17, 507–517. https://doi.org/10.1162/0898929053279522

Getzmann, S., Golob, E.J., Wascher, E., 2016. Focused and divided attention in a simulated cocktail-party situation: ERP evidence from younger and older adults. Neurobiology of Aging 41, 138–149. https://doi.org/10.1016/j.neurobiolaging.2016.02.018

Getzmann, S., Lewald, J., 2010. Effects of natural versus artificial spatial cues on electrophysiological correlates of auditory motion. Hearing Research. 259, 44–54. https://doi.org/10.1016/j.heares.2009.09.021

Giuliano, R.J., Karns, C.M., Neville, H.J., Hillyard, S.A., 2014. Early Auditory Evoked Potential Is Modulated by Selective Attention and Related to Individual Differences in Visual Working Memory Capacity. Journal of Cognitive Neuroscience. 26(12), 2682–2690. https://doi.org/10.1162/jocn_a_00684

Glyde, H., Buchholz, J.M., Dillon, H., Cameron, S., Hickson, L., 2013. The importance of interaural time differences and level differences in spatial release from masking. Journal of the Acoustical Society of America. 134, EL147-EL152. https://doi.org/10.1121/1.4812441

Golomb, J.D., Chun, M.M., Mazer, J.A., 2008. The Native Coordinate System of Spatial Attention Is Retinotopic. Journal of Neuroscience. 28, 10654–10662. https://doi.org/10.1523/jneurosci.2525-08.2008

Green, J.J., Doesburg, S.M., Ward, L.M., McDonald, J.J., 2011. Electrical Neuroimaging of Voluntary Audiospatial Attention: Evidence for a Supramodal Attention Control Network. Journal of Neuroscience. 31, 3560–3564. https://doi.org/10.1523/JNEUROSCI.5758-10.2011

Haegens, S., Cousijn, H., Wallis, G., Harrison, P.J., Nobre, A.C., 2014. Inter- and intra-individual variability in alpha peak frequency. NeuroImage 92, 46–55. https://doi.org/10.1016/j.neuroimage.2014.01.049

Haegens, S., Handel, B.F., Jensen, O., 2011. Top-Down Controlled Alpha Band Activity in Somatosensory Areas Determines Behavioral Performance in a Discrimination Task. Journal of Neuroscience. 31, 5197–5204. https://doi.org/10.1523/jneurosci.5199-10.2011

Handy, T.C., Soltani, M., Mangun, G.R., 2001. Perceptual Load and Visuocortical Processing: Event-Related Potentials Reveal Sensory-Level Selection. Psychological Science. 12, 213–218. https://doi.org/10.1111/1467-9280.00338

Hartmann, W.M., Wittenberg, A., 1996. On the externalization of sound images. Journal of the Acoustical Society of America 99, 3678–3688.

He, B.J., Snyder, A.Z., Vincent, J.L., Epstein, A., Shulman, G.L., Corbetta, M., 2007. Breakdown of Functional Connectivity in Frontoparietal Networks Underlies Behavioral Deficits in Spatial Neglect. Neuron 53, 905–918. https://doi.org/10.1016/j.neuron.2007.02.013

Helfer, K.S., Freyman, R.L., 2005. The role of visual speech cues in reducing energetic and informational masking. Journal of the Acoustical Society of America. 117, 842–849. https://doi.org/10.1121/1.1836832

Helfrich, R.F., Schneider, T.R., Rach, S., Trautmann-Lengsfeld, S.A., Engel, A.K., Herrmann, C.S., 2014. Entrainment of brain oscillations by transcranial alternating current stimulation. Current Biology. 24, 333–339. https://doi.org/10.1016/j.cub.2013.12.041

Herrmann, C.S., Knight, R.T., 2001. Mechanisms of human attention: event-related potentials and oscillations. Neuroscience and Biobehavioral Reviews. 25, 465–476. https://doi.org/10.1016/S0149-7634(01)00027-6

Herrmann, C.S., Strüber, D., Helfrich, R.F., Engel, A.K., 2016. EEG oscillations: From correlation to causality. International Journal of Psychophysiology. 103, 12–21. https://doi.org/10.1016/j.ijpsycho.2015.02.003

Higgins, N.C., McLaughlin, S.A., Rinne, T., Stecker, G.C., 2017. Evidence for cue-independent spatial representation in the human auditory cortex during active listening. Proceedings of the National Academy of Sciences of the United States of America. 114, E7602--E7611. https://doi.org/10.1073/pnas.1707522114

Hillyard, S. a, Hink, R.F., Schwent, V.L., Picton, T.W., 2007. Electrical Signs of Selective Attention in the Human Brain tions of cardiac output and peripheral Electrical Signs of Selective Attention in the Human Brain. Psychophysiology 182, 177–180.

Hillyard, S.A., Anllo-Vento, L., 1998. Event-related brain potentials in the study of visual selective attention. Proceedings of the National Academy of Sciences of the United States of America. 95, 781–787. https://doi.org/10.1073/pnas.95.3.781

Hillyard, S.A., Hink, R.F., Schwent, V.L., Picton, T.W., 1973. Electrical signs of selective attention in the human brain. Science 182(4108), 177–180.

https://doi.org/10.1126/SCIENCE.182.4108.177

Hillyard, S.A., Vogel, E.K., Luck, S.J., 1998. Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. Philosophical Transactions of the Royal Society of London. B. Biological Sciences. 353, 1257–70. https://doi.org/10.1098/rstb.1998.0281

Hopfinger, J., Luck, S., Hillyard, S., 2004. Selective attention: Electrophysiological and neuromagnetic studies. In M.S. Gazziniga (ed.) The Cognitive Neurosciences (pp. 561–574). Cambridge, MA: MIT Press..

Huang, S., Chang, W.T., Belliveau, J.W., Hämäläinen, M., Ahveninen, J., 2014. Lateralized parietotemporal oscillatory phase synchronization during auditory selective attention. NeuroImage 86, 461–469. https://doi.org/10.1016/j.neuroimage.2013.10.043

Ikkai, A., Dandekar, S., Curtis, C.E., 2016. Lateralization in Alpha-Band Oscillations Predicts the Locus and Spatial Distribution of Attention. PLoS One 11, e0154796. https://doi.org/10.1371/journal.pone.0154796

Itoh, K., Yumoto, M., Uno, A., Kurauchi, T., Kaga, K., 2000. Temporal stream of cortical representation for auditory spatial localization in human hemispheres. Neuroscience Letters. 292, 215–219. https://doi.org/10.1016/S0304-3940(00)01465-8

Jacobs, J., 2013. Hippocampal theta oscillations are slower in humans than in rodents: implications for models of spatial navigation and memory. Philosophical Transactions of the Royal Society of London. B. Biological Sciences. 369, 20130304–20130304. https://doi.org/10.1098/rstb.2013.0304

Jensen, O., Mazaheri, A., 2010. Shaping Functional Architecture by Oscillatory Alpha Activity: Gating by Inhibition. Frontiers in Human Neuroscience. 4, 186. https://doi.org/10.3389/fnhum.2010.00186

Johnson, B.W., Hautus, M.J., 2010. Processing of binaural spatial information in human auditory cortex: {Neuromagnetic} responses to interaural timing and level differences. Neuropsychologia 48, 2610–2619. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2010.05.008

Kaiser, J., Lutzenberger, W., Preissl, H., Ackermann, H., Birbaumer, N., 2000. Right-Hemisphere Dominance for the Processing of Sound-Source Lateralization. Journal of Neuroscience. 20(17), 6631–6639.

Kanai, R., Chaieb, L., Antal, A., Walsh, V., Paulus, W., 2008. Frequency-Dependent Electrical Stimulation of the Visual Cortex. Current Biology. 18, 1839–1843. https://doi.org/10.1016/j.cub.2008.10.027

Kelly, S.P., Lalor, E.C., Reilly, R.B., Foxe, J.J., 2006. Increases in Alpha Oscillatory Power Reflect an Active Retinotopic Mechanism for Distracter Suppression During Sustained Visuospatial Attention. Journal of Neurophysiology. 95, 3844–3851. https://doi.org/10.1152/jn.01234.2005

Kerlin, J.R., Shahin, A.J., Miller, L.M., 2010. Attentional Gain Control of Ongoing Cortical Speech Representations in a “Cocktail Party.” Journal of Neuroscience. 30, 620–628. https://doi.org/10.1523/JNEUROSCI.3631-09.2010

Kidd, G., Arbogast, T.L., Mason, C.R., Gallun, F.J., 2005. The advantage of knowing where to listen. Journal of the Acoustical Society of America. 118, 3804–3815.

Kidd, G., Mason, C.R., Best, V., Marrone, N., Marrone, N., 2010. Stimulus factors influencing spatial release from speech-on-speech masking. Journal of the Acoustical Society of America. 128, 1965–78. https://doi.org/10.1121/1.3478781

Kidd, G., Mason, C.R., Rohtla, T.L., Deliwala, P.S., 1998. Release from masking due to spatial separation of sources in the identification of nonspeech auditory patterns. Journal of the Acoustical Society of America. 104, 422–431. https://doi.org/10.1121/1.423246

Klatt, L.-I., Getzmann, S., Wascher, E., Schneider, D., 2018. The contribution of selective spatial attention to sound detection and sound localization: Evidence from event-related potentials and lateralized alpha oscillations. Biological Psychology. 138, 133–145. https://doi.org/10.1016/J.BIOPSYCHO.2018.08.019

Klaus, J., Schutter, D.J.L.G., 2018. Non-invasive brain stimulation to investigate language production in healthy speakers: A meta-analysis. Brain and Cognition. 123, 10–22. https://doi.org/10.1016/J.BANDC.2018.02.007

Klimesch, W., 2012. Alpha-band oscillations, attention, and controlled access to stored information. Trends in Cognitive Sciences. 16, 606–617. https://doi.org/10.1016/J.TICS.2012.10.007

Klimesch, W., 2011. Evoked alpha and early access to the knowledge system: The P1 inhibition timing hypothesis. Brain Research. 1408, 52–71. https://doi.org/10.1016/j.brainres.2011.06.003

Klimesch, W., 1999. EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. Brain Research Reviews. 29, 169–195. https://doi.org/10.1016/S0165-0173(98)00056-3

Klimesch, W., 1997. EEG-alpha rhythms and memory processes. International Journal of Psychophysiology. 26, 319–340. https://doi.org/10.1016/S0167-8760(97)00773-3

Klimesch, W., Sauseng, P., Hanslmayr, S., 2007. EEG alpha oscillations: The inhibition-timing hypothesis. Brain Research Reviews. 53, 63–88. https://doi.org/10.1016/j.brainresrev.2006.06.003

Knecht, S., 2000. Handedness and hemispheric language dominance in healthy humans. Brain 123, 2512–2518. https://doi.org/10.1093/brain/123.12.2512

Koehler, S., Lauer, P., Schreppel, T., Jacob, C., Heine, M., Boreatti-Hümmer, A., Fallgatter, A.J., Herrmann, M.J., 2009. Increased EEG power density in alpha and theta bands in adult ADHD patients. Journal of Neural Transmission. 116, 97–104. https://doi.org/10.1007/s00702-008-0157-x

Kong, L., Michalka, S.W., Rosen, M.L., Sheremata, S.L., Swisher, J.D., Shinn-Cunningham, B.G., Somers, D.C., 2014. Auditory Spatial Attention Representations in the Human Cerebral Cortex. Cerebral Cortex 24, 773–784. https://doi.org/10.1093/cercor/bhs359

Kuhlman, W.N., 1978. Functional topography of the human mu rhythm. Electroencephalography and Clinical Neurophysiology. 44, 83–93. https://doi.org/10.1016/0013-4694(78)90107-4

Larson, E., Lee, A.K.C., 2014. Switching auditory attention using spatial and non-spatial features recruits different cortical networks. NeuroImage 84, 681–687. https://doi.org/10.1016/J.NEUROIMAGE.2013.09.061

Lee, A.K.C., Rajaram, S., Xia, J., Bharadwaj, H., Larson, E., Hämäläinen, M.S., Shinn-Cunningham, B.G., 2013. Auditory Selective Attention Reveals Preparatory Activity in Different Cortical Regions for Selection Based on Source Location and Source Pitch. Frontiers in Neuroscience. 6. https://doi.org/10.3389/fnins.2012.00190

Leino, S., May, P.J.C., Alku, P., Liikkanen, L.A., Tiitinen, H., 2007. The contribution of high frequencies to human brain activity underlying horizontal localization of natural spatial sounds. BMC Neuroscience. 8, 78. https://doi.org/10.1186/1471-2202-8-78

Lim, S.-J., Wöstmann, M., Obleser, J., 2015. Selective Attention to Auditory Memory Neurally Enhances Perceptual Precision. Journal of Neuroscience. 35, 16094–16104. https://doi.org/10.1523/JNEUROSCI.2674-15.2015

Loiselle, L.H., Dorman, M.F., Yost, W.A., Cook, S.J., Gifford, R.H., 2016. Using ILD or ITD Cues for Sound Source Localization and Speech Understanding in a Complex Listening Environment by Listeners With Bilateral and With Hearing-Preservation Cochlear Implants. Journal of Speech, Language, and Hearing Research. 59, 810–818. https://doi.org/10.1044/2015_JSLHR-H-14-0355

Luck, S.J., 1995. Multiple mechanisms of visual-spatial attention: recent evidence from human electrophysiology. Behavioral Brain Research. 71, 113–123. https://doi.org/10.1016/0166-4328(95)00041-0

Luck, S.J., Heinze, H.J., Mangun, G.R., Hillyard, S.A., 1990. Visual event-related potentials index focused attention within bilateral stimulus arrays. II. Functional dissociation of P1 and N1 components. Electroencephalography and Clinical Neurophysiology. 75, 528–542. https://doi.org/10.1016/0013-4694(90)90139-B

Luck, S.J., Hillyard, S.A., Mouloua, M., Woldorff, M.G., Clark, V.P., Hawkins, H.L., 1994. Effects of spatial cuing on luminance detectability: Psychophysical and electrophysiological evidence for early selection. Journal of Experimental Psychology. Human Perception and Performance. 20, 887–904. https://doi.org/10.1037/0096-1523.20.4.887

Luo, H., Liu, Z., Poeppel, D., 2010. Auditory Cortex Tracks Both Auditory and Visual Stimulus Dynamics Using Low-Frequency Neuronal Phase Modulation. PLoS Biology. 8, e1000445. https://doi.org/10.1371/journal.pbio.1000445

Macaluso, E., Frith, C.D., Driver, J., 2002. Supramodal effects of covert spatial orienting triggered by visual or tactile events. Journal of Cognitive Neuroscience. 14, 389–401. https://doi.org/10.1162/089892902317361912

Macaluso, E., Frith, C.D., Driver, J., 2000. Modulation of human visual cortex by crossmodal spatial attention. Science. 289(5482), 1206–1208. https://doi.org/10.1126/science.289.5482.1206

Madden, D.J., Plude, D.J., 1993. Selective Preservation of Selective Attention. In J. Cerella J.M. Rybash, W. Hoyer and M.L. Commons (Eds.) Adult Information Processing: Limits on Loss (pp. 273–300). San Diego, CA: Academic Press.

Maddox, R.K., Shinn-Cunningham, B.G., 2012. Influence of Task-Relevant and Task-Irrelevant Feature Continuity on Selective Auditory Attention. Journal of the Association for Research in Otolaryngology. 13, 119–129. https://doi.org/10.1007/s10162-011-0299-7

Majdak, Baumgartner, R., Jenny, C., 2018. Formation of three-dimensional auditory space. Available from https://arxiv.org/abs/1901.03990

Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. Journal of Neuroscience Methods 164, 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Marzecová, A., Schettino, A., Widmann, A., SanMiguel, I., Kotz, S.A., Schröger, E., 2018. Attentional gain is modulated by probabilistic feature expectations in a spatial cueing task: ERP evidence. Scientific Reports 8. https://doi.org/10.1038/s41598-017-18347-1

Mathewson, K.J., Jetha, M.K., Drmic, I.E., Bryson, S.E., Goldberg, J.O., Schmidt, L.A., 2012. Regional EEG alpha power, coherence, and behavioral symptomatology in autism spectrum disorder. Clinical Neurophysiology. 123, 1798–1809. https://doi.org/10.1016/J.CLINPH.2012.02.061

Matsumoto, H., Ugawa, Y., 2017. Adverse events of tDCS and tACS: A review. Clinical Neurophysiology Practice. 2, 19–25. https://doi.org/10.1016/j.cnp.2016.12.003

Mehraei, G., Shinn-Cunningham, B., Dau, T., 2018. Influence of talker discontinuity on cortical dynamics of auditory spatial attention. NeuroImage. 179, 548–556. https://doi.org/10.1016/j.neuroimage.2018.06.067

Mesgarani, N., Chang, E.F., 2012. Selective cortical representation of attended speaker in multi-talker speech perception. Nature 485, 233–236. https://doi.org/10.1038/nature11020

Mesulam, M.-M., 1999. Spatial attention and neglect: parietal, frontal and cingulate contributions to the mental representation and attentional targeting of salient extrapersonal events. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences. 354, 1325–1346. https://doi.org/10.1098/rstb.1999.0482

Middlebrooks, J.C., 2015. Sound localization. Handbook of Clinical Neurology, The Human Auditory System 129, 99–116. https://doi.org/10.1016/B978-0-444-62630-1.00006-8

Moliadze, V., Atalay, D., Antal, A., Paulus, W., 2012. Close to threshold transcranial electrical stimulation preferentially activates inhibitory networks before switching to excitation with higher intensities. Brain Stimulation. 5, 505–511. https://doi.org/10.1016/j.brs.2011.11.004

Mondor, T.A., Bregman, A.S., 1994. Allocating attention to frequency regions. Perception and Psychophysics. 56, 268–276. https://doi.org/10.3758/BF03209761

Murray, M.M., Brunet, D., Michel, C.M., 2008. Topographic ERP Analyses: A Step-by-Step Tutorial Review. Brain Topography. 20, 249–264. https://doi.org/10.1007/s10548-008-0054-5

Näätänen, R., Sams, M., Alho, K., Paavilainen, P., Reinikainen, K., Sokolov, E.N., 1988. Frequency and location specificify of the human vertex N1 wave. Electroencephalography and Clinical Neurophysiology. 69, 523–531. https://doi.org/10.1016/0013-4694(88)90164-2

Neuling, T., Rach, S., Herrmann, C.S., 2013. Orchestrating neuronal networks: Sustained after-effects of transcranial alternating current stimulation depend upon brain states. Frontiers in Human Neuroscience. 7, 161. https://doi.org/10.3389/fnhum.2013.00161

Nguyen, J., Deng, Y., Reinhart, R.M.G., 2018. Brain-state determines learning improvements after transcranial alternating-current stimulation to frontal cortex. Brain Stimulation. 11(4), 723–726. https://doi.org/10.1016/j.brs.2018.02.008

Nitsche, M.A., Cohen, L.G., Wassermann, E.M., Priori, A., Lang, N., Antal, A., Paulus, W., Hummel, F., Boggio, P.S., Fregni, F., Pascual-Leone, A., 2008. Transcranial direct current stimulation: State of the art 2008. Brain Stimulation. 1(3), 206–223. https://doi.org/10.1016/j.brs.2008.06.004

Noble, W., Perrett, S., 2002. Hearing speech against spatially separate competing speech versus competing noise. Perception & Psychophysics. 64, 1325–1336. https://doi.org/10.3758/BF03194775

Okazaki, Y.O., Horschig, J.M., Luther, L., Oostenveld, R., Murakami, I., Jensen, O., 2015. Real-time MEG neurofeedback training of posterior alpha activity modulates subsequent visual detection performance. NeuroImage 107, 323–332. https://doi.org/10.1016/j.neuroimage.2014.12.014

Oldfield, R.C., 1971. The assessment and analysis of handedness: The Edinburgh inventory. Neuropsychologia 9, 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.M., 2011. FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Computational Intelligence and Neuroscience. 2011, 156869. https://doi.org/10.1155/2011/156869

Ozen, S., Sirota, A., Belluscio, M.A., Anastassiou, C.A., Stark, E., Koch, C., Buzsaki, G., 2010. Transcranial Electric Stimulation Entrains Cortical Neuronal Populations in Rats. Journal of Neuroscience. 30, 11476–11485. https://doi.org/10.1523/JNEUROSCI.5252-09.2010

Palomäki, K.J., Tiitinen, H., Mäkinen, V., May, P.J.C., Alku, P., 2005. Spatial processing in human auditory cortex: The effects of 3D, ITD, and ILD stimulation techniques. Cognitive Brain Research. 24, 364–379. https://doi.org/10.1016/j.cogbrainres.2005.02.013

Pavani, F., Ládavas, E., Driver, J., 2003. Auditory and multisensory aspects of visuospatial neglect. Trends in Cognitive Sciences. 7, 407–414. https://doi.org/10.1016/S1364-6613(03)00189-X

Pfurtscheller, G., Neuper, C., Krausz, G., 2000. Functional dissociation of lower and upper frequency mu rhythms in relation to voluntary limb movement. Clinical Neurophysiology. 111, 1873–1879.

Plack, C.J., Barker, D., 2014. Pitch coding and pitch processing in the human brain. Hearing Research. 307, 53–64. https://doi.org/10.1016/J.HEARES.2013.07.020

Posner, M.I., Snyder, C.R., Davidson, B.J., 1980. Attention and the detection of signals. Journal of Experimental Psychology. General. 109, 160–174. https://doi.org/10.1037/0096-3445.109.2.160

Pouget, A., Driver, J., 2000. Relating unilateral neglect to the neural coding of space. Current Opinion in Neurobiology. 10(2), 242–249. https://doi.org/10.1016/S0959-4388(00)00077-5

Rauschecker, J.P., Tian, B., 2000. Mechanisms and streams for processing of “what” and “where” in auditory cortex. Proceedings of the National Academy of Sciences of the United States of America. 97, 11800–11806. https://doi.org/10.1073/pnas.97.22.11800

Razali, N., Wah, Y.B. 2011. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. Journal of Statistical Modeling and Analytics. 2(1), 21–33.

Reato, D., Rahman, A., Bikson, M., Parra, L.C., 2013. Effects of weak transcranial alternating current stimulation on brain activity—a review of known mechanisms from animal studies. Frontiers in Human Neuroscience. 7, 1–8. https://doi.org/10.3389/fnhum.2013.00687

Reato, D., Rahman, A., Bikson, M., Parra, L.C., 2010. Low-Intensity Electrical Stimulation Affects Network Dynamics by Modulating Population Rate and Spike Timing. Journal of Neuroscience. 30, 15067–15079. https://doi.org/10.1523/JNEUROSCI.2059-10.2010

Reinhart, R.M.G., Cosman, J.D., Fukuda, K., Woodman, G.F., 2017. Using transcranial direct-current stimulation (tDCS) to understand cognitive processing. Attention, Perception, & Psychophysics. 79, 3–23. https://doi.org/10.3758/s13414-016-1224-2

Rihs, T.A., Michel, C.M., Thut, G., 2007a. Mechanisms of selective inhibition in visual spatial attention are indexed by α-band EEG synchronization. European Journal of Neuroscience. 25, 603–610. https://doi.org/10.1111/j.1460-9568.2007.05278.x

Rimmele, J.M., Zion Golumbic, E., Schröger, E., Poeppel, D., 2015. The effects of selective attention and speech acoustics on neural speech-tracking in a multi-talker scene. Cortex 68, 144–154. https://doi.org/10.1016/J.CORTEX.2014.12.014

Roberts, K.L., Hall, D.A., 2008. Examining a Supramodal Network for Conflict Processing: A Systematic Review and Novel Functional Magnetic Resonance Imaging Data for Related Visual and Auditory Stroop Tasks. Journal of Cognitive. Neuroscience. 20, 1063–1078. https://doi.org/10.1162/jocn.2008.20074

Röder, B., Teder-Sälejärvi, W., Sterr, A., Rösler, F., Hillyard, S.A., Neville, H.J., 1999. Improved auditory spatial tuning in blind humans. Nature 400, 162–166. https://doi.org/10.1038/22106

Romei, V., Gross, J., Thut, G., 2010. On the Role of Prestimulus Alpha Rhythms over Occipito-Parietal Areas in Visual Input Regulation: Correlation or Causation? Journal of Neuroscience. 30, 8692–8697. https://doi.org/10.1523/jneurosci.0160-10.2010

Sach, A.J., Hill, N.I., Bailey, P.J., 2000. Auditory spatial attention using interaural time differences. Journal of Experimental Psychology. Human Perception and Performance. 26, 717–729. https://doi.org/10.1037/0096-1523.26.2.717

Salmi, J., Rinne, T., Koistinen, S., Salonen, O., Alho, K., 2009. Brain networks of bottom-up triggered and top-down controlled shifting of auditory attention. Brain Research. 1286, 155–164. https://doi.org/10.1016/j.brainres.2009.06.083

Salminen, N.H., Takanen, M., Santala, O., Lamminsalo, J., Altoè, A., Pulkki, V., 2015. Integrated processing of spatial cues in human auditory cortex. Hearing Research. 327, 143–152. https://doi.org/10.1016/J.HEARES.2015.06.006

Samaha, J., Bauer, P., Cimaroli, S., Postle, B.R., 2015. Top-down control of the phase of alpha-band oscillations as a mechanism for temporal prediction. Proceedings of the National Academy of Sciences of the United States of America. 112, 8439–8444. https://doi.org/10.1073/pnas.1503686112

Samaha, J., Sprague, T.C., Postle, B.R., 2016b. Decoding and Reconstructing the Focus of Spatial Attention from the Topography of Alpha-band Oscillations. Journal of Cognitive Neuroscience. 28, 1090–1097. https://doi.org/10.1162/jocn_a_00955

Sauseng, P., Klimesch, W., Doppelmayr, M., Pecherstorfer, T., Freunberger, R., Hanslmayr, S., 2005. EEG alpha synchronization and functional coupling during top-down processing in a working memory task. Human Brain Mapping. 26, 148–155. https://doi.org/10.1002/hbm.20150

Sauseng, P., Klimesch, W., Stadler, W., Schabus, M., Doppelmayr, M., Hanslmayr, S., Gruber, W.R., Birbaumer, N., 2005. A shift of visual spatial attention is selectively associated with human EEG alpha activity. European Journal of Neuroscience. 22, 2917–2926. https://doi.org/10.1111/j.1460-9568.2005.04482.x

Schonwiesner, M., Krumbholz, K., Rubsamen, R., Fink, G.R., von Cramon, D.Y., 2006. Hemispheric Asymmetry for Auditory Processing in the Human Auditory Brain Stem, Thalamus, and Cortex. Cerebral Cortex 17, 492–499. https://doi.org/10.1093/cercor/bhj165

Schreuder, M., Rost, T., Tangermann, M., 2011. Listen, You are Writing! Speeding up Online Spelling with a Dynamic Auditory BCI. Frontiers in Neuroscience. 5, 112. https://doi.org/10.3389/fnins.2011.00112

Schröger, E., 1996. Interaural time and level differences: Integrated or separated processing? Hearing Research. 96, 191–198. https://doi.org/10.1016/0378-5955(96)00066-4

Schwartz, A., McDermott, J.H., Shinn-Cunningham, B., 2012. Spatial cues alone produce inaccurate sound segregation: The effect of interaural time differences. Journal of the Acoustical Society of America. 132, 357–368. https://doi.org/10.1121/1.4718637

Serences, J.T., Yantis, S., 2006. Selective visual attention and perceptual coherence. Trends in Cognitive Sciences. 10, 38–45. https://doi.org/10.1016/J.TICS.2005.11.008

Serences, J.T., Yantis, S., Culberson, A., Awh, E., 2004. Preparatory Activity in Visual Cortex Indexes Distractor Suppression During Covert Spatial Orienting. Journal of Neurophysiology. 92, 3538–3545. https://doi.org/10.1152/jn.00435.2004

Shinn-Cunningham, B., Best, V., Lee, A.K.C., 2017. Auditory Object Formation and Selection. In J. Middlebrooks, J.Z. Simon, A.N. Popper, and R.R. Fay (eds.) The Auditory System at the Cocktail Party (pp. 7–40). Springer Cham. https://doi.org/10.1007/978-3-319-51662-2_2

Shinn-Cunningham, B.G., 2008. Object-based auditory and visual attention. Trends in Cognitive. Sciences. 12, 182–186. https://doi.org/10.1016/j.tics.2008.02.003

Shinn-Cunningham, B.G., Kopco, N., Martin, T.J., 2005. Localizing nearby sound sources in a classroom: Binaural room impulse responses. Journal of the Acoustical Society of America. 117, 3100–3115. https://doi.org/10.1121/1.1872572

Shulman, G.L., Pope, D.L.W., Astafiev, S. V., McAvoy, M.P., Snyder, A.Z., Corbetta, M., 2010. Right Hemisphere Dominance during Spatial Selective Attention and Target Detection Occurs Outside the Dorsal Frontoparietal Network. Journal of Neuroscience. 30, 3640–3651. https://doi.org/10.1523/jneurosci.4085-09.2010

Siegel, M., Donner, T.H., Engel, A.K., 2012. Spectral fingerprints of large-scale neuronal interactions. Nature Reviews. Neuroscience. 13, 121–134. https://doi.org/10.1038/nrn3137

Silver, M.A., Ress, D., Heeger, D.J., 2005. Topographic maps of visual spatial attention in human parietal cortex. Journal of Neurophysiology. 94, 1358–1371. https://doi.org/10.1152/jn.01316.2004

Skrandies, W., 1990. Global field power and topographic similarity. Brain Topography. 3, 137–141.

Slagter, H.A., Prinssen, S., Reteig, L.C., Mazaheri, A., 2016. Facilitation and inhibition in attention: Functional dissociation of pre-stimulus alpha activity, P1, and N1 components. NeuroImage 125, 25–35. https://doi.org/10.1016/j.neuroimage.2015.09.058

Smith, R.C.G., Price, S.R., 2014. Modelling of Human Low Frequency Sound Localization Acuity Demonstrates Dominance of Spatial Variation of Interaural Time Difference and Suggests Uniform Just-Noticeable Differences in Interaural Time Difference. PLoS One 9, e89033. https://doi.org/10.1371/journal.pone.0089033

Snyder, A.C., Foxe, J.J., 2010. Anticipatory Attentional Suppression of Visual Features Indexed by Oscillatory Alpha-Band Power Increases:A High-Density Electrical Mapping Study. Journal of Neuroscience. 30, 4024–4032. https://doi.org/10.1523/JNEUROSCI.5684-09.2010

Snyder, J.S., Large, E.W., 2005. Gamma-band activity reflects the metric structure of rhythmic tone sequences. Cognitive Brain Research. 24, 117–126. https://doi.org/10.1016/j.cogbrainres.2004.12.014

Stevens, C., Fanning, J., Coch, D., Sanders, L., Neville, H., 2008. Neural mechanisms of selective auditory attention are enhanced by computerized training: Electrophysiological evidence from language-impaired and typically developing children. Brain Research 1205, 55–69. https://doi.org/10.1016/J.BRAINRES.2007.10.108

Strauß, A., Wöstmann, M., Obleser, J., 2014. Cortical alpha oscillations as a tool for auditory selective inhibition. Frontiers in Human Neuroscience. 8, 1–7. https://doi.org/10.3389/fnhum.2014.00350

Strauss, D.J., Francis, A.L., 2017. Toward a taxonomic model of attention in effortful listening. Cognitive. Affective & Behavioral Neuroscience. 17, 809–825. https://doi.org/10.3758/s13415-017-0513-0

Strüber, D., Rach, S., Neuling, T., Herrmann, C.S., 2015. On the possible role of stimulation duration for after-effects of transcranial alternating current stimulation. Frontiers in Cellular Neuroscience. 9, 311. https://doi.org/10.3389/fncel.2015.00311

Swisher, J.D., Halko, M.A., Merabet, L.B., McMains, S.A., Somers, D.C., 2007. Visual Topography of Human Intraparietal Sulcus. Journal of Neuroscience. 27, 5326–5337. https://doi.org/10.1523/JNEUROSCI.0991-07.2007

Tardif, E., Murray, M.M., Meylan, R., Spierer, L., Clarke, S., 2006. The spatio-temporal brain dynamics of processing and integrating sound localization cues in humans. Brain Research. 1092, 161–176. https://doi.org/10.1016/J.BRAINRES.2006.03.095

Tavakoli, A. V., Yun, K., 2017. Transcranial Alternating Current Stimulation (tACS) Mechanisms and Protocols. Frontiers in Cellular Neuroscience. 11, 1–10. https://doi.org/10.3389/fncel.2017.00214

Teder-Sälejärvi, W.A., Hillyard, S.A., 1998. The gradient of spatial auditory attention in free field: An event-related potential study. Perception & Psychophysics. 60, 1228–1242. https://doi.org/10.3758/BF03206172

Teshiba, T.M., Ling, J., Ruhl, D.A., Bedrick, B.S., Peña, A., Mayer, A.R., 2013. Evoked and Intrinsic Asymmetries during Auditory Attention: Implications for the Contralateral and Neglect Models of Functioning. Cerebral Cortex 23, 560–569. https://doi.org/10.1093/cercor/bhs039

Thut, G., Nietzel A, Brandt SA, P.-L.A., 2006. Alpha-Band Electroencephalographic Activity over Occipital Cortex Indexes Visuospatial Attention Bias and Predicts Visual Target Detection. Journal of Neuroscience. 26(37), 9494–9502. https://doi.org/10.1523/JNEUROSCI.0875-06.2006

Thut, G., Veniero, D., Romei, V., Miniussi, C., Schyns, P., Gross, J., 2011. Rhythmic TMS causes local entrainment of natural oscillatory signatures. Current Biology. 21, 1176–1185. https://doi.org/10.1016/j.cub.2011.05.049

Triesman, A.M., 1969. Strategies and models of selective attention. Psychological Review. 76, 282–299.

Truong, D.Q., Magerowski, G., Blackburn, G.L., Bikson, M., Alonso-Alonso, M., 2013. Computational modeling of transcranial direct current stimulation (tDCS) in obesity: Impact of head fat and dose guidelines. NeuroImage. Clinical. 2, 759–766. https://doi.org/10.1016/j.nicl.2013.05.011

Ungan, P., Yagcioglu, S., Goksoy, C., 2001. Differences between the N1 waves of the responses to interaural time and intensity disparities: scalp topography and dipole sources. Clinical Neurophysiology. 112, 485–498. https://doi.org/10.1016/S1388-2457(00)00550-2

van Diepen, R.M., Miller, L.M., Mazaheri, A., Geng, J.J., 2016. The Role of Alpha Activity in Spatial and Feature-Based Attention. eNeuro 3, e0204–16.2016.

https://doi.org/10.1523/ENEURO.0204-16.2016

van Gerven, M., Jensen, O., 2009. Attention modulations of posterior alpha as a control signal for two-dimensional brain–computer interfaces. Journal of Neuroscience Methods 179, 78–84. https://doi.org/10.1016/J.JNEUMETH.2009.01.016

Van Leeuwen, W.S., Wieneke, G., Spoelstra, P., Versteeg, H., 1978. Lack of bilateral coherence of mu rhythm. Electroencephalogr. Clinical Neurophysiology. 44, 140–146. https://doi.org/10.1016/0013-4694(78)90260-2

Veniero, D., Benwell, C.S.Y., Ahrens, M.M., Thut, G., 2017. Inconsistent effects of parietal α-tACS on Pseudoneglect across two experiments: A failed internal replication. Frontiers in Psychology. 8, 952. https://doi.org/10.3389/fpsyg.2017.00952

Villamar, M.F., Volz, M.S., Bikson, M., Datta, A., DaSilva, A.F., Fregni, F., 2013. Technique and Considerations in the Use of 4x1 Ring High-definition Transcranial Direct Current Stimulation (HD-tDCS). Journal of Visualized Experiments: JoVE. (77), e50309. https://doi.org/10.3791/50309

Vossen, A., Gross, J., Thut, G., 2015. Alpha power increase after transcranial alternating current stimulation at alpha frequency (a-tACS) reflects plastic changes rather than entrainment. Brain Stimulation. 8, 499–508. https://doi.org/10.1016/j.brs.2014.12.004

Ward, L.M., 1994. Supramodal and modality-specific mechanisms for stimulus-driven shifts of auditory and visual attention. Canadian Journal of Experimental Psychology. 48, 242–259. https://doi.org/10.1037/1196-1961.48.2.242

Weisz, N., Hartmann, T., Müller, N., Lorenz, I., Obleser, J., 2011. Alpha Rhythms in Audition: Cognitive and Clinical Perspectives. Frontiers in Psychology. 2. https://doi.org/10.3389/fpsyg.2011.00073

Weisz, N., Obleser, J., 2014. Synchronisation signatures in the listening brain: A perspective from non-invasive neuroelectrophysiology. Hearing Research. 307, 16–28. https://doi.org/10.1016/j.heares.2013.07.009

Wightman, F.L., Kistler, D.J., 1992. The dominant role of low-frequency interaural time differences in sound localization. Journal of the Acoustical Society of America. 91, 1648–1661. https://doi.org/10.1121/1.402445

Wilcoxon, F., 1945. Individual Comparisons by Ranking Methods. Biometrics Bulletin. 1(6), 80. https://doi.org/10.2307/3001968

Wild-Wall, N., Falkenstein, M., 2010. Age-dependent impairment of auditory processing under spatially focused and divided attention: An electrophysiological study. Biological Psychology. 83, 27–36. https://doi.org/10.1016/j.biopsycho.2009.09.011

Willems, R.M., Der Haegen, L. Van, Fisher, S.E., Francks, C., 2014. On the other hand: Including left-handers in cognitive neuroscience and neurogenetics. Nature Reviews. Neuroscience. 15, 193–201. https://doi.org/10.1038/nrn3679

Woldorff, M., Gallen, C., et al., 1993. 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. 90(18), 8722–8726.

Woldorff, M., Hillyard, S.A., 1991. Modulation of early auditory processing during selective listening to rapidly presented tones. Electroencephalography and Clinical Neurophysiology. 79(3), 170–191.

Worden, M.S., Foxe, J.J., Wang, N., Simpson, G. V., 2000. Anticipatory Biasing of Visuospatial Attention Indexed by Retinotopically Specific α-Band Electroencephalography Increases over Occipital Cortex. Journal of Neuroscience. 20(6), RC63. https://doi.org/10.1523/jneurosci.20-06-j0002.2000

Wöstmann, M., Herrmann, B., Maess, B., Obleser, J., 2016. Spatiotemporal dynamics of auditory attention synchronize with speech. Proceedings of the National Academy of Sciences of the United States of America. 113, 3873–3878. https://doi.org/10.1073/pnas.1523357113

Wu, C.T., Weissman, D.H., Roberts, K.C., Woldorff, M.G., 2007. The neural circuitry underlying the executive control of auditory spatial attention. Brain Research. 1134, 187–198. https://doi.org/10.1016/j.brainres.2006.11.088

Wyart, V., Dehaene, S., Tallon-Baudry, C., 2012. Early dissociation between neural signatures of endogenous spatial attention and perceptual awareness during visual masking. Frontiers in Human Neuroscience. 6, 16. https://doi.org/10.3389/fnhum.2012.00016

CURRICULUM VITAE

[redacted]

[redacted]

[redacted]

[redacted]