

Nonspatial Features Reduce the Reliance on Sustained Spatial Auditory Attention

Lia M. Bonacci,¹ Scott Bressler,¹ and Barbara G. Shinn-Cunningham^{1,2}

Objective: Top-down spatial attention is effective at selecting a target sound from a mixture. However, nonspatial features often distinguish sources in addition to location. This study explores whether redundant nonspatial features are used to maintain selective auditory attention for a spatially defined target.

Design: We recorded electroencephalography while subjects focused attention on one of three simultaneous melodies. In one experiment, subjects ($n = 17$) were given an auditory cue indicating both the location and pitch of the target melody. In a second experiment ($n = 17$ subjects), the cue only indicated target location, and we compared two conditions: one in which the pitch separation of competing melodies was large, and one in which this separation was small.

Results: In both experiments, responses evoked by onsets of events in sound streams were modulated by attention, and we found no significant difference in this modulation between small and large pitch separation conditions. Therefore, the evoked response reflected that target stimuli were the focus of attention, and distractors were suppressed successfully for all experimental conditions. In all cases, parietal alpha was lateralized following the cue, but before melody onset, indicating that subjects initially focused attention in space. During the stimulus presentation, this lateralization disappeared when pitch cues were strong but remained significant when pitch cues were weak, suggesting that strong pitch cues reduced reliance on sustained spatial attention.

Conclusions: These results demonstrate that once a well-defined target stream at a known location is selected, top-down spatial attention plays a weak role in filtering out a segregated competing stream.

Key words: Auditory attention, Auditory event-related potentials, Electroencephalography, Nonspatial attention, Parietal alpha oscillations, Spatial attention.

(Ear & Hearing 2020;XX;00–00)

INTRODUCTION

Spatial features of an auditory object are often useful for focusing attention in noisy environments—if the spatial location of the object is known, then that information can be used to select this target in one location while suppressing irrelevant objects in another (Shinn-Cunningham 2008). Often, however, additional features, such as pitch, differentiate target from distractor streams. It is therefore unclear to what extent spatial features are used when listeners must maintain attention on an auditory stream if other features also differentiate competing streams.

Previous work has shown that both spatial and nonspatial features interact to guide a listener's attention to an ongoing stream. In particular, discontinuity in the spatial location of a stream has been shown to disrupt attention when nonspatial features are otherwise continuous (Best et al. 2008; Maddox &

Shinn-Cunningham 2012; Kreitewolf et al. 2018; Mehraei et al. 2018; Deng et al. 2019). While these results suggest that both spatial and nonspatial features contribute to how an object is formed and selected bottom-up, other studies have shown that spatial and nonspatial features are used differentially when directing attention top-down, depending on the current goal (Lee et al. 2012; Maddox & Shinn-Cunningham 2012; Larson & Lee 2014; Deng et al. 2019). However, it is still unclear to what degree volitional, top-down attention biases selection based on spatial features when redundant nonspatial features also differentiate target from distractors. Neural correlates of top-down attention obtained from noninvasive electroencephalography (EEG) may provide insights into the strategies that listeners use when sound sources have multiple distinguishing features.

Selective auditory attention modulates the amplitude of event-related potentials (ERPs) in auditory cortex measured using EEG; ERPs evoked by one stream are greater when that stream is attended compared with when it is ignored (Choi et al. 2013, 2014). Selective attention can be deployed based on a target sound's location or based on nonspatial features, such as pitch and timbre (Lee et al. 2012; Maddox & Shinn-Cunningham 2012; Larson & Lee 2014). Therefore, this enhancement of ERPs to attended stimuli may reflect top-down control based on any one or a combination of these features.

Spatially focused selective attention also induces changes in the distribution of parietal alpha (8 to 14 Hz) oscillatory power. Specifically, during spatial attention, alpha power increases over parietal sensors ipsilateral to the attended location (Worden et al. 2000; Banerjee et al. 2011; Foxe & Snyder 2011). This alpha lateralization has been studied extensively during visual spatial attention but has been explored to a lesser degree during auditory spatial attention (but see, e.g., Banerjee et al. 2011; Wöstmann et al. 2016; Tune et al. 2018). As noted earlier, spatial attention may not be necessary to maintain attention on a target stream once it is selected based on its location. The dynamics of alpha power lateralization can thus provide insight into whether sustained attention relies on spatial processing.

Knowing to what extent spatial features are used for top-down attention not only contributes to a better understanding of how we communicate in noisy environments, but may also provide insight for helping those who struggle to do so. Communication in these settings is particularly difficult for listeners with hearing loss, even with current assistive technology (Marrone et al. 2008; Shinn-Cunningham & Best 2008). There is therefore an increasing interest in using noninvasive EEG to predict what an individual intends to listen to in order to enhance sound at that attentional focus (Choi et al. 2013; O'Sullivan et al. 2017; Van Eyndhoven et al. 2017). However, many of these efforts rely on knowing which auditory streams are present in the scene a priori, which if integrated into assistive technology would require preliminary signal processing to automatically extract

¹Department of Biomedical Engineering, Boston University, Boston, Massachusetts, USA; and ²Neuroscience Institute, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA.

these streams. An alternative approach is to predict *where* rather than *what* the focus of attention is based on the spatial distribution of EEG alpha power. This approach could reduce the amount of preliminary processing needed to extract streams or eliminate this requirement altogether.

Yet if nonspatial cues are more informative than the available spatial cues, then individuals may depend more on these features to direct attention top-down. Therefore, if parietal alpha reflects the use of spatial features during top-down attention, then its modulation may be weak during tasks in which spatial features are redundant with other nonspatial features, calling into question the utility of parietal EEG alpha for predicting the focus of attention in real-world listening environments.

To address these questions, we measured EEG during two experiments in which subjects attended one of three competing auditory streams. Tasks were identical across experiments, but different cues were used to inform subjects as to which stream to attend. In the first experiment, an auditory cue was given that identified both the spatial location and the pitch of the target stream. Here, we asked whether subjects would orient attention in space even if they knew the pitch of the to-be-attended stream. We hypothesized that lateralization of alpha might be weak throughout attention to the cued stream because subjects did not have to orient attention in space to successfully perform the task. In the second experiment, the auditory cue only identified the spatial location of the target so that subjects would have to initially orient attention in space. We tested two conditions, presented in different blocks: one in which the pitch separation of competing melodies was large, and one in which this separation was small. We hypothesized that sustained alpha lateralization would be weak when the pitch separation was large, reflecting the fact that strong pitch cues may also be used to maintain attention to the distinct target stream, but that it would remain strong throughout trials in which spatial information was more critical for differentiating the competing streams.

MATERIALS AND METHODS

Experimental Task and Stimuli

We conducted two separate experiments, each with the same auditory selective attention task, based on that used in (Choi et al. 2014) (Fig. 1A). Three isochronous melodies were presented simultaneously from different directions—left, right, and center—using interaural time differences (ITDs) of -100 , $+100$, and 0 μsec , respectively. Previous work from our laboratory has shown that listeners can perceive clear spatial differences with these ITDs (Dai et al. 2018).

The center melody, consisting of three 1-sec notes, came on first and was always ignored. The left melody came on 0.6 sec later and consisted of four 0.6-sec notes. The right melody came on 0.15 sec after the left melody and consisted of three 0.75-sec notes. In such an arrangement, the onsets of notes in each melody were staggered in time, allowing ERPs associated with notes in each melody to be temporally isolated. In addition, the 0.15-sec lag between the first and second note onsets is just beyond the edge of how long it takes listeners to be able to recognize and orient to the second one. This makes the exogenous draw of the first onset challenging to ignore and makes the second stream harder to focus on—this was a feature of the original task design to make it more difficult (Choi et al. 2014). Note that while these distinct timings could theoretically be

used for stream segregation, they were the same across all experimental conditions and, therefore, do not contribute to any of the differences we observed between conditions.

In addition to being spatially separate and temporally staggered, the three melodies were separated by pitch differences, as indicated by Figure 1B. Notes in each melody were composed of six harmonics added in cosine phase with magnitudes inversely proportional to frequency. Melodies were composed of two notes: a high note (H) and a low note (L). These notes were arranged to form pitch contours that were “rising,” “falling,” or “zigzagging.” “Rising” melodies started on the low note and transitioned at a randomly selected point to the high note (e.g., L-L-H-H). “Falling” melodies started on the high note and transitioned at some onset to the low note (e.g., H-L-L-L). “Zigzagging” melodies started on either the high or low note, transitioned to the opposite note, and then returned to the starting note (e.g., L-H-L or H-L-H). In “zigzagging” melodies, the second pitch change always occurred between the last two notes to ensure subjects had to maintain focused attention for the duration of the auditory stream. Contours were selected independently for left, right, and center melodies, with each contour having a 1/3 chance of being chosen.

At the beginning of each trial, subjects were given an auditory cue directing them to attend either the left or the right melody. After attending the target melody, subjects had to report its pitch contour via button press. In addition to active attention trials, passive trials were included in which subjects were given a visual cue, signaling they could ignore stimuli and were to withhold a response. All cues were 100% valid. Visual feedback was given at the end of each trial to indicate if the melody was correctly identified.

Performance on active attention trials was measured as percent correct response; passive trials were counted as correct if subjects did not make a button press. We did not measure reaction time, as subjects had to withhold responses until the response period, which began 500 msec after the last stimulus ended. This allowed us to reduce motor planning and electromyogram artifacts in the EEG measures but rendered response times unreliable as a behavioral metric.

Subjects performed the experiment in front of a liquid-crystal display monitor in a sound-treated booth. Stimuli were generated using MATLAB (MathWorks, Natick, MA) with the PsychToolbox 3 extension (Brainard 1997). Sound stimuli were presented diotically via Etymotic ER-1 insert headphones (Etymotic, Elk Grove Village, IL) connected to Tucker-Davis Technologies System 3 (TDT, Alachua, FL) hardware which interfaced with MATLAB software that controlled the experiment. During the task, subjects were instructed to keep their eyes open and to foveate on a central fixation dot.

Experiment 1 • In Experiment 1, the auditory cue was a six-harmonic complex tone that was presented with the same ITD as the target melody. The fundamental frequency of this cue was also in the same pitch range as the notes composing the target. As mentioned earlier, each melody was presented in a different pitch range, as shown in Figure 1B. Within each pitch range, two of three possible fundamental frequencies were randomly selected to compose the high and low note for each two-note melody. The construction of the two-note melody from three possible fundamentals was intended to reduce predictability of the melody pattern from the first note, as originally designed by (Choi et al. 2014). The three possible fundamentals were separated by 1.65 semitones. The center melody, which was always ignored, had notes with fundamentals in the 320 to 387 Hz range. On a given trial, either the right or left

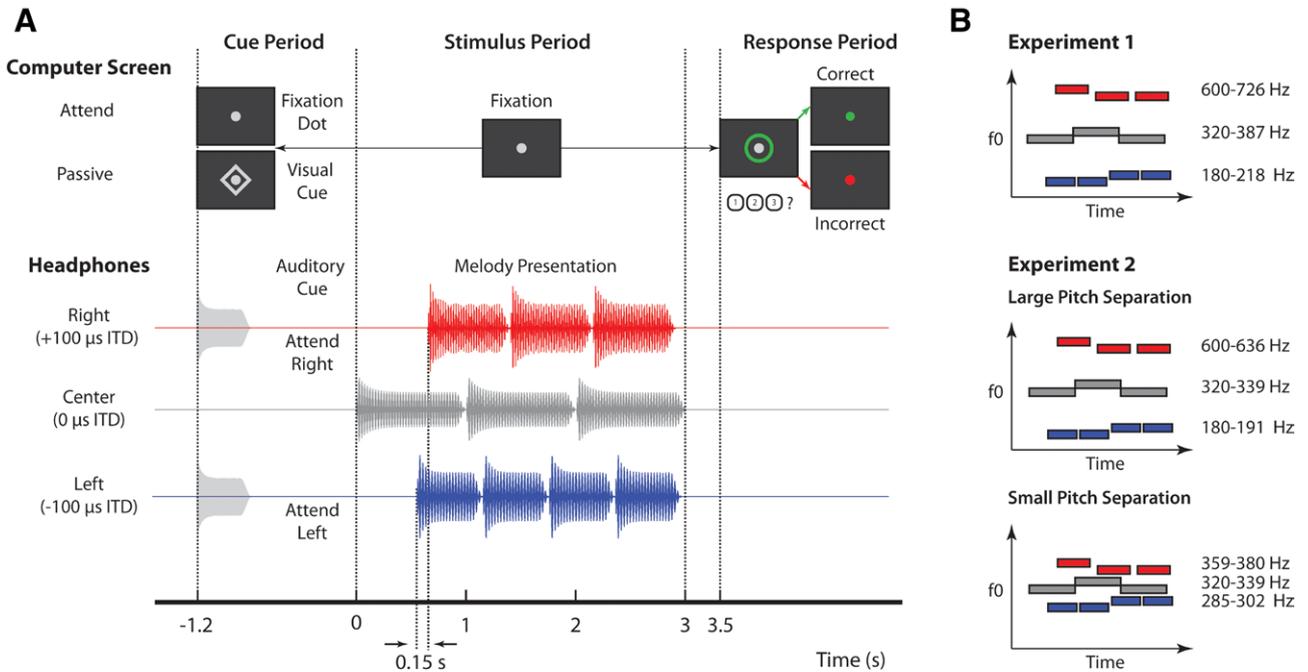


Fig. 1. Experimental design. A, Trial structure for the auditory selective attention task. For attend-left and attend-right trials, an auditory cue was presented via headphones with the same interaural time difference (ITD) as the target melody. For passive trials, a diamond appeared around a central fixation dot on screen. During the stimulus period, subjects kept their gaze on the fixation dot, while melodies were presented diotically. A green circle appeared around the fixation dot to prompt a response. Subjects were to press 1, 2, or 3 on the keyboard to indicate if the melody was “rising,” “falling,” or “zigzagging,” respectively. Visual feedback was given after button press to indicate if the target was correctly identified; for passive trials, no button press was considered a correct response. B, Left (blue), right (red), and center (gray) melodies were composed of notes with different fundamental frequencies (F0). Note that in this example, right melodies had the highest fundamentals while left melodies had the lowest fundamentals, but the opposite also occurred with equal probability. The center melody always had the same F0s, which were between F0s of the left and right melodies. Individual melodies also changed pitch over time, such that they were rising, falling, or zigzagging, illustrated by blue, red, and gray bars in this example, respectively.

melody was selected, with equal probability, to have fundamentals in the 180 to 218 Hz range. The remaining melody was selected to have fundamentals in the 600 to 726 Hz range. This pitch separation among melodies ensured that each was perceptually segregated from the others nearly universally, based on ranges defined as “always streamed” in van Noorden (1975).

Trials were arranged in 9 blocks of 30, with each block containing 1/3 attend-left and 1/3 attend-right trials presented in random order. The remaining trials were passive control trials. This resulted in 90 trials for each condition. Before performing the task, subjects were required to pass a training demo in which they were presented with a series of single melodies and asked to identify their pitch contours. Passive trials were also included in the training demo to ensure subjects knew when to withhold a response, and trials were counted as correct if no button press was made. To continue the study, subjects had to answer correctly on 10 of 12 demo trials (4 passive trials, 8 active attention trials). This requirement was included to ensure that subjects’ performance on the task was not limited by their ability to identify pitch contours but by their ability to direct attention.

Experiment 2 • In Experiment 2, the auditory cue was a white noise burst that was presented with the same ITD as the target melody. This required subjects to at least initially orient attention in space because no pitch information was available in the cue. As in Experiment 1, each melody was presented in a different pitch range. Within each pitch range, the same two fundamentals were used to compose the high and low note of each two-note melody. Note that this differed from Experiment 1,

which constructed two-note melodies from three possible fundamentals to reduce predictability of the target. For simplicity, we chose to only use the same two fundamentals, because in our experience, it does not make a substantial difference in target predictability. Within each melody, the two fundamentals composing high and low notes were separated by 1 semitone. In all trials, the center melody always had fundamentals in the middle, 320 to 339 Hz range. As in experiment 1, high and low pitch ranges were randomly assigned to the left and right melodies.

The fundamental frequency of melodies in these pitch ranges depended on the experimental block, which were one of two conditions: one in which the pitch separation of competing melodies was large and one in which it was small (Fig. 1B). In the large pitch separation condition, the low pitch melodies had fundamentals in the 180 to 191 Hz range, while the high pitch melodies had fundamentals in the 600 to 636 Hz range, creating clearly segregated streams. As in Experiment 1, this very large pitch separation among melodies should facilitate streaming nearly universally, based on the findings of van Noorden (1975). In the small pitch separation condition, fundamentals of low (285 to 302 Hz) and high (359 to 380 Hz) pitch ranges were shifted closer to that of the center melody. The resulting sound mixture was thus more difficult to automatically segregate by pitch alone. Large and small pitch separation blocks were grouped together in pairs, but the order of conditions was random for each pair of blocks (e.g., Lg-Sm-Sm-Lg-Sm-Lg-Lg-Sm).

Trials were arranged in 16 blocks of 30, with each block containing 2/5 attend-left, 2/5 attend-right, and 1/5 passive trials.

This resulted in 96 attend-left and attend-right trials in each pitch separation condition and 96 passive trials across all pitch separation conditions. After the first 8 blocks, subjects were instructed to take a break before starting the remaining set of 8 blocks. As in Experiment 1, subjects were required to pass a training demo in which they had to identify the pitch contour of a single melody presented alone. Two training blocks were given, one each for stimuli in the two pitch separation conditions. Each block contained 15 trials (3 passive trials, 12 active attention trials), and subjects had to answer correctly on 13 trials for each block to continue in the experiment.

Subjects

Data from a total of 42 subjects with normal hearing and no known neurological disorders were analyzed as part of this study—22 for Experiment 1 and 20 for Experiment 2. However, data from five subjects in Experiment 1 and three subjects in Experiment 2 had to be discarded due to too many incorrect response trials or too many trials with noisy EEG. Therefore, the final analyses shown here were performed on a total of 34 subjects—17 from Experiment 1 (8 male, mean age = 21.88, SD = 2.78) and 17 different subjects from Experiment 2 (9 male, mean age = 22.35, SD = 3.67). An audiogram was conducted for each subject to confirm that thresholds were below 20 dB HL at octave frequencies from 250 Hz to 8 kHz. Some subjects recruited for Experiment 2 were dismissed early from the study: one had audiometric thresholds above the required level, two could not give a clean EEG signal, and six failed the training demo described earlier. These subjects were compensated for their time but did not have EEG recorded. All subjects gave written informed consent before participation and were compensated at an hourly rate (\$25/hr for Experiment 1, \$15/hr for Experiment 2) as well as with a bonus for each correct response (\$0.02 per response, up to \$7.50/hr). All procedures were approved by the Boston University Institutional Review Board.

Subjects who participated in Experiment 1 also participated in an analogous visual task—not described here—during the same experimental session. Of these subjects, 12 participated in the visual task after the auditory task was complete. The remaining 5 subjects completed the visual task blocks first. Subjects who participated in Experiment 2 were not exposed to any visual analog of the task. While it is possible that the subjects who completed visual experiment before starting the auditory experiment were biased toward using spatial features for selection, we found no evidence that alpha modulation was statistically different between the subjects who completed the visual task first and those who did not.

Data Collection

EEG data were recorded at a sampling rate of 2048 Hz using the BioSemi ActiveTwo system and its ActiveView acquisition software (BioSemi, Amsterdam, The Netherlands). A 64-channel cap with electrode positions arranged according to the international 10–20 system was used for measurement, along with two reference electrodes placed on the mastoids. An additional three electrodes were placed around the eyes for electrooculogram measurement, which was included with EEG data for the purpose of removing eye blinks with independent component analysis (ICA). Event triggers were driven by MATLAB software running the experimental task and generated by

Tucker-Davis Technologies System 3 (TDT, Alachua, FL) hardware that interfaced with the computer recording EEG data. In Experiment 2, RME Fireface UXC hardware was used instead of the TDT for trigger generation. An EyeLink Plus 1000 (SR Research, Ottawa, Ontario, Canada) eye tracker was used in Experiment 2 to ensure that subjects did not close or move their eyes during the task. In Experiment 1, subjects were instructed to fixate on a central fixation dot, but eye tracking was not recorded or monitored during the experiment.

Data Analysis

EEG Processing • The EEGLAB toolbox for MATLAB (Delorme & Makeig 2004) was used to process raw EEG data. Raw EEG data were first rereferenced to the average between two mastoid electrodes and downsampled to 256 Hz. A finite impulse response zero-phase filter with cutoffs at 1 and 20 Hz was then applied to the signal. Eye blinks were removed using ICA in EEGLAB. ICA was performed on the combined set of electrooculogram and EEG sensors. Components that were visually identified as containing only eye blinks were flagged manually and removed from EEG by the software, as described in Jung et al. (2000) and Chaumon et al. (2015). Epochs with amplitudes over $\pm 100 \mu\text{V}$ were rejected along with trials in which subjects gave an incorrect response. We chose to discard incorrect trials because we were interested primarily in differences in strategy for performing the task; this requires EEG measures that accurately reflect successful task performance. EEG data from subjects who had fewer than 60 correct trials were discarded before further analysis. CSD Toolbox (Kayser & Tenke 2006) was used to transform EEG data from voltage to current source density. This technique was employed to reduce spatially correlated EEG noise, which is desirable when localizing parietal alpha power across the scalp (Kayser & Tenke 2015; McFarland 2015).

Event-Related Potential • Because magnitude estimates of time-domain waveforms (such as ERPs) are often affected greatly by outliers, the ERP time course was estimated using a bootstrap procedure. To reduce the likelihood of outliers unduly influencing the results, the average time course was first calculated across 100 randomly chosen trials with replacement within a single subject and condition. This procedure was repeated 200 times, and each subject's estimated ERP was taken as the average across these 200 trials. Each subject's ERPs were then normalized by dividing the entire time series by the average amplitude of the N1 response to the first center melody note onset, averaged across all trials and channels. This step ensured that all ERPs were similar in magnitude across subjects. The first center melody note onset was selected because it was previously shown to elicit a strong N1 response that is not modulated by attention (Choi et al. 2014), presumably due to the salience of the initial sound onset eliciting involuntary attention. Grand averages were obtained for each condition by averaging the normalized ERP amplitudes across subjects.

N1 amplitudes were extracted for each subject from the normalized ERP time courses. These ERPs were first averaged across 17 frontocentral channels where responses were largest (Fp1, Fp2, AF3, AF7, AF4, AF8, F1, F3, F5, F7, F2, F4, F6, F8, Fpz, AFz, Fz). This normalized channel average was then averaged across subjects to estimate the N1 timings for each note onset. These times, selected based on the largest negative value of the ERP in a window between 75 and 240 msec following

each stimulus onset, were then used to estimate N1 amplitudes for each subject's channel-averaged ERP. The ERP was averaged in a 50-msec window centered around each of the selected time points to quantify N1 amplitude in response to each note. Each subject's ERP was visually inspected to ensure that N1s were correctly identified.

Attentional modulation of the N1 was quantified for each subject using an attentional modulation index (AMI_{N1}) given by Eq. 1.

$$AMI_{N1} = N1_{\text{attend}} - N1_{\text{ignore}} \quad (1)$$

Here, $N1_{\text{attend}}$ is the negative of the ERP amplitude elicited by the onset of a particular note at the determined N1 time when that note was attended; $N1_{\text{ignore}}$ is the negative of the ERP amplitude elicited at the same time when this note was ignored. Note that ERP amplitudes were multiplied by -1 first so that large positive values of AMI_{N1} indicate that N1 amplitudes were larger when notes were attended, relative to when they were ignored, as expected a priori. AMI_{N1} was calculated for each note in both left and right melodies and averaged to quantify overall modulation of the N1. The N1 to the first leading left onset was not included in this average because it has previously been shown in similar paradigms to elicit a strong automatic response regardless of cue condition. This decision was made a priori based on previous findings from our laboratory with a similar experimental paradigm (see Choi et al. 2014; Bressler et al. 2017).

Induced Alpha Power • To obtain the induced alpha response, it was necessary to first remove phase-locked activity. To achieve this, the phase-locked or evoked response (ERP) was calculated as the mean across trials and subtracted from each trial, leaving only the induced, non-phase-locked activity. Power at each frequency in the alpha band (8 to 14 Hz) was estimated for each trial using a short-time Fourier transform. An individual alpha frequency was selected for each subject by finding the frequency in the range of 8 to 14 Hz with the greatest magnitude across attend-left and attend-right conditions in 20 parietal and occipital channels (P2, P4, P6, P8, P10, PO4, PO8, O2, P1, P3, P5, P7, P9, PO3, PO7, O1, Pz, POz, Oz, Iz). Power was extracted at this individual alpha frequency to produce a single time series for each trial in each EEG channel. To reduce the likelihood of outliers influencing each subject's average alpha power estimates, the bootstrap procedure described earlier for estimating the ERP was used to estimate each subject's average induced alpha power for each experimental condition. Normalization was performed on individual subject trial-averaged time series by dividing each time point by that subject's average alpha power across time, sensors, and experimental condition. Grand averages were obtained by averaging these normalized time series across subjects. Quantities shown on topoplots represent averages across the cue period (-1.2 to 0 sec) or stimulus period (0 to 2.44 sec).

An AMI of alpha power, AMI_{α} , was also calculated for each subject. Calculation of AMI_{α} is given by Eq. 2.

$$AMI_{\alpha} = \frac{\alpha_{\text{ipsi}} - \alpha_{\text{contra}}}{\alpha_{\text{ipsi}} + \alpha_{\text{contra}}} \quad (2)$$

In Eq. 2, α_{ipsi} is the average alpha power during the stimulus or cue period, measured ipsilateral to the cued sequence, or rather contralateral to the ignored sequence; α_{contra} is this average alpha power, measured contralateral to the cued sequence.

Large positive values of AMI_{α} indicate that alpha power was overall larger ipsilateral to cued stimuli (i.e., the alpha response was larger over cortices processing ignored information). Averages were calculated across left and right parietal and occipital channels separately, depending on the attention condition (i.e., left channels for α_{ipsi} in attend-left trials and right channels for α_{ipsi} in attend-right trials). These averages were then collapsed across attention conditions and parietal sensors to quantify α_{ipsi} and α_{contra} .

Significance Testing • Statistical testing was performed solely on the final AMI_{N1} and AMI_{α} values calculated for each subject. However, we first provide some descriptive results to allow readers to build insight into how these values reflect attentional modulation during the task.

For Experiment 1, we performed statistical testing to determine if modulation of N1 was significantly greater than zero. For this purpose, we used a one-sample, one-sided t test on AMI_{N1} data. We also wanted to determine if alpha lateralization, indexed by AMI_{α} , was significantly greater than zero in both the cue and stimulus periods. Again we used a one-sample, one-sided t test. The same statistical procedures were used in Experiment 2 to determine if AMI_{N1} and AMI_{α} were significantly greater than zero. We also hypothesized that AMI_{α} would be greater in the small pitch separation condition than in the large pitch separation condition. To determine if this difference was significant, we performed paired-sample, one-sided t tests for values measured during cue and stimulus periods. Multiple comparisons procedures were performed before determining significance. AMI_{N1} was also compared between large and small pitch separation conditions using a paired-sample t test. To correct for multiple comparisons, the Bonferroni-Holm procedure was used. For all comparisons, we also computed Cohen's d effect sizes.

Passive Trial Analysis • We performed a comparison of EEG measures between active attention and passive trials for Experiment 2 to better understand how each quantity reflects enhancement and suppression during selective attention. Because there were fewer passive trials relative to active attention trials in each block, passive trials were averaged across all blocks for a fair comparison with attend-left and attend-right trials.

To better understand how attentional modulation of the N1 response reflects enhancement of target stimuli and suppression of ignored stimuli, we compared N1 amplitudes between active attention trials and passive trials. We calculated Attention Enhancement as the difference of the N1 amplitude between attend and passive conditions and Ignore Suppression as the difference between passive and ignore conditions. These values were averaged for each subject across all N1 onsets (excluding the first leading onset). Positive values of Attention Enhancement indicate that attention to a stimulus enhances the N1 responses it evokes. Positive values of Ignore Suppression indicate that attention leads to active suppression of irrelevant stimuli. We used one-tailed t tests to determine if Attention Enhancement and Ignore Suppression were significantly greater than zero. We also used paired t tests to determine if Attention Enhancement was significantly greater than Ignore Suppression, or vice versa. The Bonferroni-Holm procedure was used to correct for multiple comparisons, and effect sizes were calculated as Cohen's d .

We also compared alpha power between active attention and passive trials. For a given set of parietal sensors, we calculated

Contralateral Suppression during ignore conditions as the difference in alpha power between attend-ipsilateral (i.e., ignore-contralateral) and passive trials. If alpha is an active suppression mechanism, this quantity should be significantly greater than zero. For a given set of sensors, we also calculated Contralateral Suppression during attend conditions as the difference in alpha power between attend-contralateral and passive trials. We used one-tailed t tests to determine if Contralateral Suppression was significantly greater than zero for both attend and ignore conditions. Correction for multiple comparisons was again performed using the Bonferroni-Holm procedure, and effect sizes were calculated as Cohen's d .

RESULTS

Behavior

Differences in Performance Existed Between Attend-Left and Attend-Right Trials in Experiment 2 • Performance, measured as percent correct response, is displayed for both experiments in Figure 2. Overall, subjects performed well above chance, suggesting successful focus of attention. In Experiment 1, no significant differences were found between attend-left and attend-right trials ($p = 0.24$, paired t test; $d = 0.29$). Differences in performance were found in Experiment 2, however. A two-way repeated-measures analysis of variance found a significant interaction between pitch condition and cue condition ($F(1,16) = 22.35$, $p < 0.001$). In the large pitch separation condition, subjects performed better on attend-right trials than on attend-left trials ($p = 0.015$, paired t test, corrected for 4 comparisons; $d = 0.79$). The opposite was true for the small pitch separation condition ($p = 0.026$, paired t test, corrected for 4 comparisons; $d = 0.68$). In comparing spatial attention conditions across pitch separation conditions, there was no significant difference in performance on attend-left trials between large and small pitch separation conditions ($p = 0.94$, paired t test, corrected for 4 comparisons; $d = 0.020$). For attend-right trials, however, performance was significantly greater in the large pitch separation condition ($p < 0.001$, paired t test, corrected for 4 comparisons; $d = 1.20$).

These performance differences may be explained by differences in the bottom-up salience of the melodies in the two conditions. Recall that the right melody always lagged the left melody in time. Therefore, in the large pitch separation condition, even though the leading (left) melody may have captured

attention first, the right melody had a distinctive pitch that caused the lagging melody to be heard as a new event automatically. In the small pitch separation condition, the lagging melody had a similar pitch to the leading melody, which likely made the melody onset less clear and salient.

Event-Related Potential

In Both Experiments, the N1 Response Was Similarly Modulated by Selective Attention • In Experiment 1, N1 amplitudes were modulated by attention (Fig. 3A, top). Specifically, N1 amplitudes were more negative in response to left note onsets (blue vertical lines) when those notes were attended (blue trace) compared with when they were ignored (attend-right trials, red trace). Similarly, N1 amplitudes were more negative in response to right note onsets (red vertical lines) when those notes were attended (red trace) compared with when they were ignored (attend-left trials, blue trace). The same modulation of the N1 was observed in Experiment 2, both in the large pitch separation condition (Fig. 3A, middle) and the small pitch separation condition (Fig. 3A, bottom). This modulation was quantified using the AMI described in Eq. 1. In both experiments, AMI_{N1} was significantly greater than zero ($p < 0.001$, t test, corrected for 2 comparisons for Experiment 2), and effect sizes were large ($d = 1.19$, $d = 1.61$, $d = 1.60$, for Experiment 1, and large and small pitch separation conditions of Experiment 2, respectively), indicating that the N1s were larger in response to attended stimuli than ignored stimuli in all experimental conditions. AMI_{N1} was also compared between pitch separation conditions in Experiment 2, but no significant difference in modulation was found, and the observed effect was small ($p = 0.26$, t test; $d = 0.29$). Thus, the degree of N1 modulation did not change significantly based on the degree of pitch information available in this experiment, suggesting that subjects selected target stimuli regardless of the available pitch cues.

Induced Alpha Power

In examining the time course of alpha power, we found that differences among experimental conditions could not be resolved at particularly fine time scales, likely due to the noisy nature of alpha power estimates. In addition, the relatively narrow bandwidth of analysis limits the rate at which estimated power can change. Therefore, the time course of alpha power was not

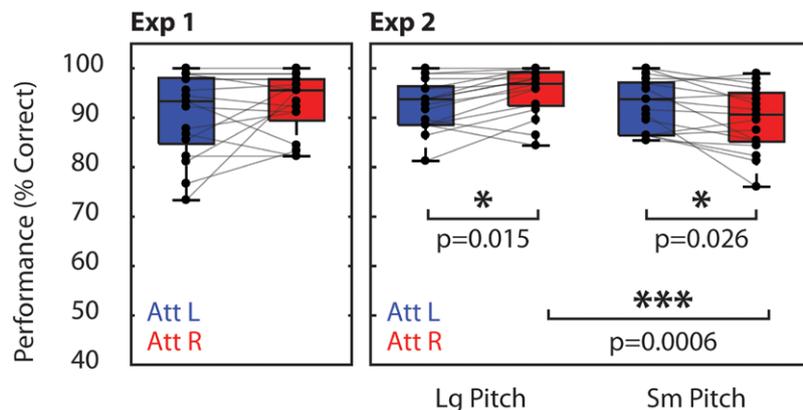


Fig. 2. Percent correct scores for Experiments 1 (left) and 2 (right). Asterisks indicate significant differences between conditions (paired t test). All displayed p values were corrected for multiple comparisons using the Bonferroni-Holm procedure.

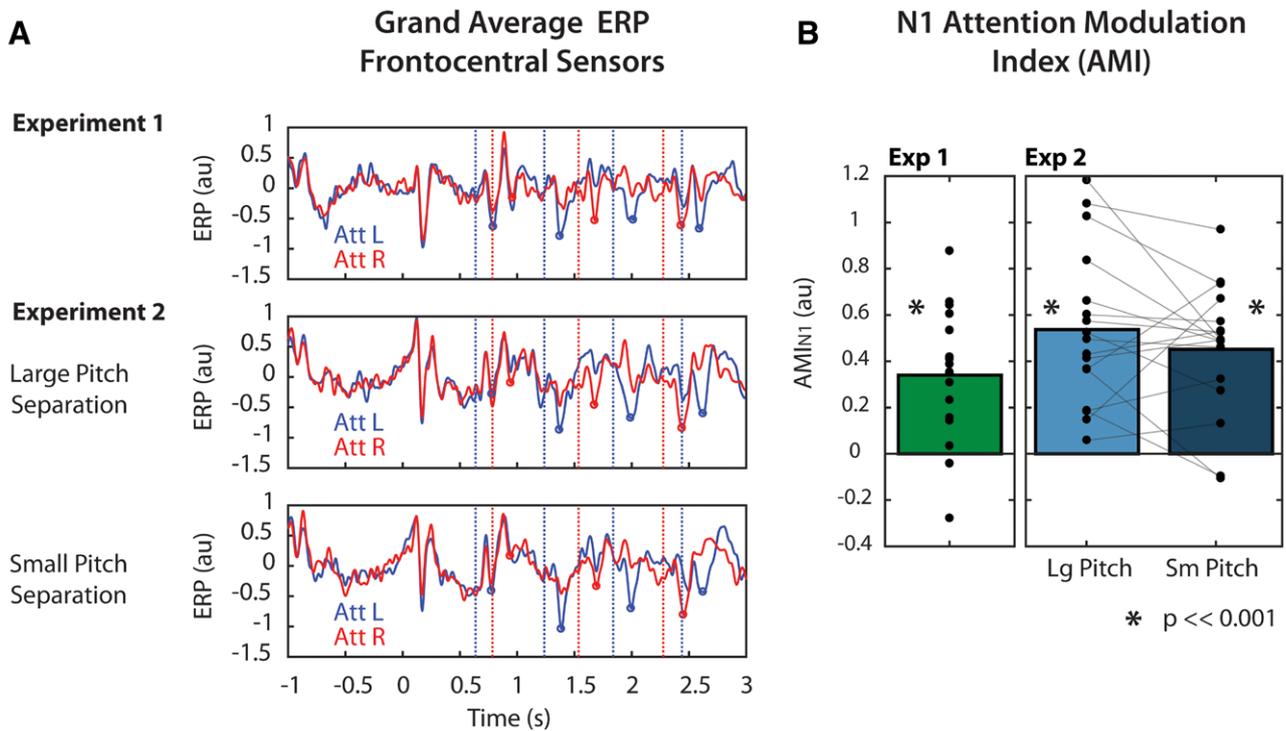


Fig. 3. Attentional modulation of the N1 response. A, Grand average ($n = 17$) normalized event-related potential (ERP) responses over time in Experiment 1 (top) and Experiment 2 (bottom). ERPs were averaged across frontocentral electroencephalography (EEG) sensors. Red and blue vertical lines indicate right and left note onset times, respectively. Red and blue circles indicate the identified N1 peak amplitudes in response to right and left notes, respectively. B, N1 modulation summarized as AMI_{N1} . Individual points indicate individual subject AMI_{N1} , calculated from Eq. 1. Asterisks indicate that AMI_{N1} was significantly greater than zero at the $p = 0.001$ significance level (one-sided, one-sample t test, corrected for multiple comparisons). AMI indicates attentional modulation index.

particularly informative. However, after averaging across time in both the cue ($t = -1.2$ to 0 sec) and stimulus ($t = 0$ to 2.44 sec) periods, we observed clear differences among experimental conditions.

During the Cue Period, Alpha Power Was Lateralized Across Parietal Sensors in All Experimental Conditions • Grand average alpha power differences, averaged over the cue period, are shown in the left panels of Figure 4A and 4B for both experiments. Figure 4A shows alpha power differences between attend-left and attend-right trials. In all experimental conditions, average alpha power was greater in left parietal sensors during attend-left trials than during attend-right trials. Similarly, in right parietal sensors, alpha power was greater in attend-right trials than in attend-left trials. Figure 4B shows these differences collapsed across left and right parietal sensors, so that alpha is represented as the difference between ipsilateral and contralateral attention conditions. During the cue period, alpha power was greater when attended stimuli were ipsilateral to a given parietal sensor than when attended stimuli were contralateral to that same sensor. This suggests that alpha increased contralateral to the ignored location, supporting the idea that alpha reflects suppression of distractors.

Figure 4C shows attention modulation indices (AMI_{α}), which are based on the ipsilateral/contralateral differences shown in Figure 4B. In Experiment 1, alpha was lateralized during the cue period (green bars), and this lateralization was significantly greater than zero ($p = 0.018$, t test, corrected for 2 comparisons; $d = 0.64$). In Experiment 2, AMI_{α} was also measured during the cue period and was significantly greater than zero for both the large (light blue bars) and small (dark blue

bars) pitch separation conditions ($p = 0.006, 0.029$, respectively, t test, corrected for 4 comparisons; $d = 0.82, d = 0.58$).

We also wondered if AMI_{α} during the cue period was different between small and large pitch separation conditions in Experiment 2 but found no significant difference ($p = 0.51$, paired t test, corrected for 2 comparisons; $d = 0.0074$). These results suggest that subjects initially oriented attention using known spatial features of the target.

During the Stimulus Period, Alpha Lateralization Was Weak When Pitch Cues Were Strong • Grand average alpha power differences, averaged over the stimulus period, are shown in the right panels of Figure 4A and 4B for both experiments. While alpha power was lateralized in both experiments during the cue period, this lateralization only persisted strongly during the stimulus period in the small pitch separation condition of Experiment 2. Here, alpha power was larger in left parietal sensors during attend-left trials and larger in right parietal sensors during attend-right trials. In the large pitch separation condition, alpha power was larger in left parietal sensors during attend-left trials. In right parietal sensors, alpha was also greater during attend-left trials, but this difference was smaller than in left parietal sensors. In Experiment 1, alpha power in right parietal sensors was greater during attend-right trials. In left parietal sensors, there was not a large difference between attend-left and attend-right trials.

Figure 4B shows these differences collapsed across parietal sensors. Here, we see that there was not a large overall difference in alpha lateralization between ipsilateral and contralateral attention trials during the stimulus period in Experiment 1 or in the large pitch separation condition of Experiment 2. In the small pitch separation condition, however, the difference between

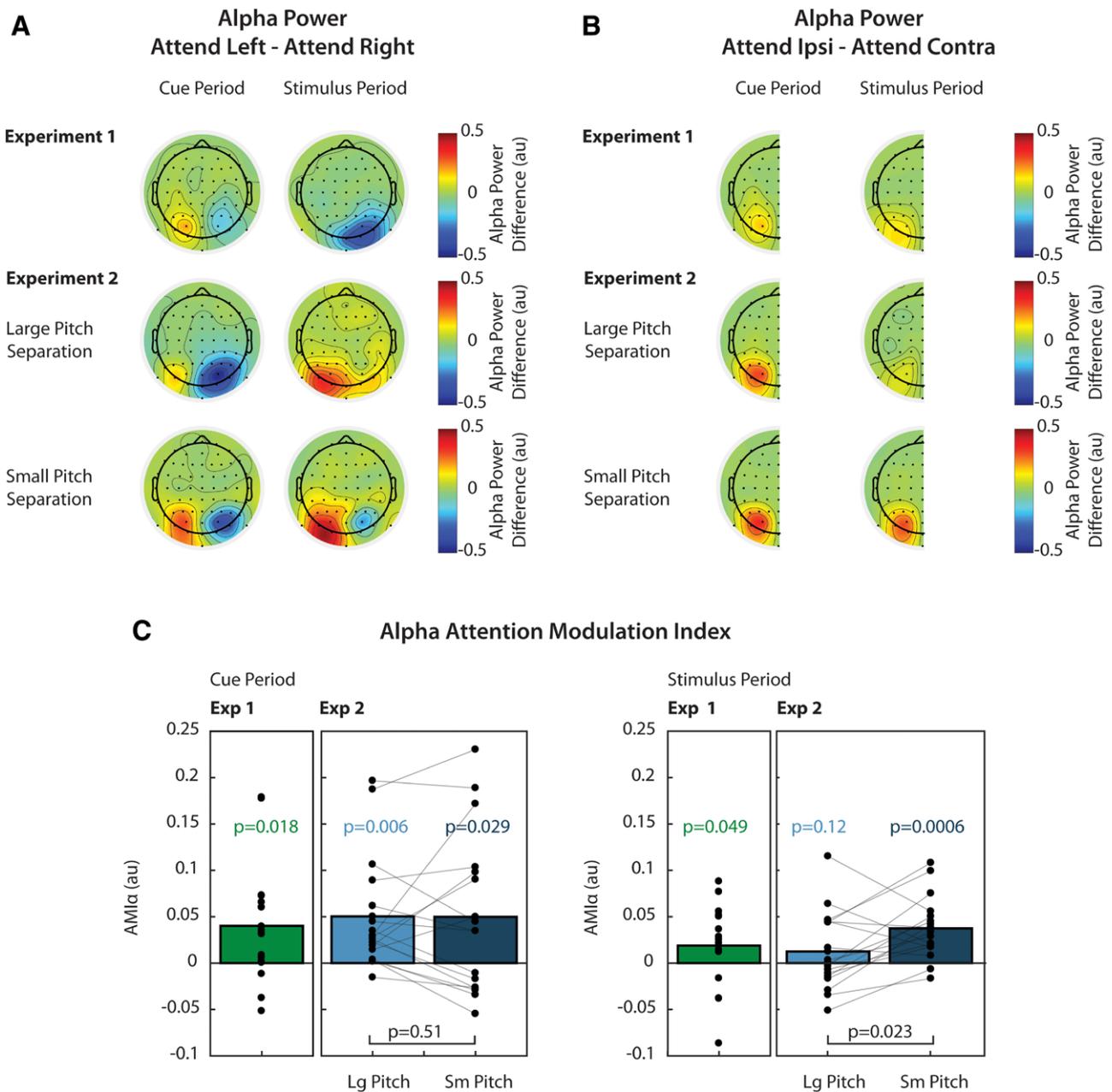


Fig. 4. Attentional modulation of alpha power. A, Grand average ($n = 17$) normalized alpha power differences between attend-left and attend-right trials. For each channel, alpha power was averaged across time during the cue period (left, $t = -1.2$ to 0 sec) or during the stimulus period (right, $t = 0$ to 2.44 sec). B, Grand average alpha power differences between ipsilateral and contralateral attention conditions, collapsed across left and right parietal channels. C, AMI_{α} , calculated from Eq. 2, during the cue period (left) and stimulus period (right). p values over individual bars are the result of t tests used to determine if was significantly greater than zero. Comparisons between conditions in Experiment 2 are also shown by brackets and associated p values. All displayed p values were corrected for multiple comparisons. AMI indicates attentional modulation index.

alpha power in ipsilateral and contralateral attention trials was similar to that observed during the cue period. These differences are represented as AMI_{α} in Figure 4C. While AMI_{α} was significantly greater than zero in Experiment 1 at the $\alpha = 0.05$ significance level ($p = 0.049$, t test, corrected for 2 comparisons), the observed effect size was small ($d = 0.43$) and was likely driven by modulation in right parietal sensors (see Fig. 4A). In the large pitch separation condition of Experiment 2, AMI_{α} was not significantly greater than zero, and the effect size was also small ($p = 0.12$, t test, corrected for 4 comparisons; $d = 0.296$). In the small pitch separation condition, however, AMI_{α} was

significantly greater than zero, and the effect was large ($p < 0.001$, t test, corrected for 4 comparisons; $d = 1.12$). We also determined that AMI_{α} was significantly larger in the small pitch separation condition compared with the large pitch separation condition ($p = 0.023$, paired t test, corrected for 2 comparisons; $d = 0.61$). **AMI_{α} Was Not Correlated With Performance Measures or AMI_{NI}** • We wished to determine if the degree of attentional suppression, as measured by AMI_{α} , correlated with performance scores, which differed between small and large pitch separation conditions. We asked this question to determine if the larger AMI_{α} observed in the small pitch separation

condition could be explained by task difficulty. We therefore looked for correlations between AMI_{α} measures and percent correct scores. For alpha power, we calculated AMI_{α} separately for left and right parietal channels and looked for correlations with percent correct scores in attend-left or attend-right trials (4 comparisons in each pitch separation condition). We found no significant correlation between any combination of AMI_{α} and percent correct scores (Spearman rank correlation, $p > 0.2$ for all comparisons).

We also wondered if the degree of alpha modulation, measured by AMI_{α} , was correlated with the degree of N1 modulation, measured by AMI_{N1} in any of the experimental conditions. However, we found no such correlation (Spearman rank correlation, $p > 0.2$ for all comparisons). This result is likely due to the fact that individual subject measures of alpha power are noisy, and effects are only observed at the group level.

Passive Trials

Grand average ERPs from Experiment 2, with the passive-trial average superimposed over active attention trials, are shown in Figure 5A. N1 modulation in the active trials appears to buildup over time relative to the passive condition. In general, passive N1 amplitudes appear to fall in between those of the two active conditions: responses evoked by a given note are smaller in magnitude in the passive condition than when that note is attended (red/blue trace at red/blue dotted lines, respectively) but larger in magnitude than when the note is ignored (red/blue trace at blue/red dotted lines, respectively). This suggests that the N1 reflects both enhancement of attended stimuli and suppression of ignored stimuli.

These differences in N1 amplitude between active and passive trials were quantified as Attention Enhancement (i.e., the difference between attend and passive) and Ignore Suppression (i.e., the difference between passive and ignore), as shown in Figure 5B. We found that for both pitch separation conditions in Experiment 2, these metrics of enhancement and suppression are both significantly greater than zero ($p < 0.01$, corrected for 4 comparisons). Effect sizes were large for both Attention Enhancement ($d = 1.209$, $d = 1.276$, large and small pitch separations, respectively) and Ignore Suppression ($d = 0.918$, $d = 0.933$, large and small pitch separations, respectively). However, within each condition, we found no significant difference between the strength of enhancement and suppression ($p > 0.09$), and effect sizes were small for both pitch separation conditions ($d = 0.441$, $d = 0.223$, large and small pitch separations, respectively). Thus, we did not find evidence that attention acts more strongly through either enhancement or suppression, but instead causes both.

The effects of attention on the strength of parietal alpha oscillations, relative to the passive condition, are shown in Figure 5C. Because alpha modulation was overall weaker during the stimulus period, as shown in Figure 4, this analysis was only performed during the cue period. Left topoplots show the difference in alpha power between attend-left and passive conditions, averaged over the cue period. In left and right parietal sensors, alpha power is greater in the attend-left condition than in the passive condition, supporting the idea that alpha modulation reflects suppression of ignored stimuli. Due to parietal asymmetry, objects in the right hemifield are represented in both left and right parietal sensors. Therefore, when attending left stimuli

(i.e., ignoring right stimuli), alpha increases with respect to passive in both left and right parietal sensors. The topoplots on the right show the difference in alpha power between attend-right and passive conditions. Again, alpha power is greater in the attend-right condition, and positive values of this difference are concentrated over right parietal sensors, which represent stimuli in the ignored, left hemifield.

Differences in alpha power relative to passive were quantified as Contralateral Suppression for both attend and ignore conditions (Attend Contralateral–Passive or Ignore Contralateral–Passive), as shown in Figure 5D. Due to asymmetry in the parietal representation of space (see Fig. 5C), these quantities were calculated separately for left and right parietal sensors. In right parietal sensors, Contralateral Suppression for the ignore condition was significantly greater than zero ($p < 0.05$, corrected for 8 comparisons) for both large ($d = 0.999$) and small ($d = 0.836$) pitch separation conditions, suggesting that alpha power modulation reflects active suppression of irrelevant stimuli. Contralateral Suppression for the attend condition was not significantly greater than zero ($p > 0.08$, for both comparisons), however. If alpha modulation is a suppression mechanism, this result is consistent, because target stimuli should not be suppressed. The fact that this metric appears to be greater than zero at all ($d = 0.507$, $d = 0.598$, large and small pitch separations, respectively) is likely due to the observed parietal asymmetry. In left parietal sensors, no measure of Contralateral Suppression was significantly greater than zero ($p > 0.08$ for all comparisons). While we expected this quantity to be significantly greater than zero in the ignore condition, this was not the case; however, we observed small to medium effect sizes ($d = 0.443$, $d = 0.591$, large and small pitch separations, respectively). This is again likely due to the parietal asymmetry—suppression may be distributed across both parietal hemifields, rather than concentrated over just one. As expected, however, Contralateral Suppression was not significantly greater than zero in the attend condition, and observed effects were relatively small ($d = 0.371$, $d = 0.196$ for large and small pitch separations, respectively). Together, the results in Figure 5D hint that alpha oscillations act to suppress contralateral stimuli, though the parietal asymmetry makes these findings difficult to interpret.

DISCUSSION

Modulation of the N1 Response Reflects Selection of Target Stimuli, but Does Not Suggest Which Feature Was Used to Perform Selection

In both experiments, we observed similar modulation of the N1 response in frontocentral channels, suggesting that there was no effect of pitch cue strength on N1 modulation. Modulation of N1s from auditory cortex reflects enhancement of target stimuli as well as suppression of distracting stimuli, as shown here in Figure 5A and 5B, as well as in previous studies (Choi et al. 2014; Kong et al. 2014). Therefore, the fact that we observed no difference in N1 modulation between experimental conditions suggests that subjects were able to focus attention on the target stream, even when pitch differences were small. A previous study that used a similar paradigm (Choi et al. 2014) found that when competing melodies were in overlapping pitch ranges, N1 modulation was degraded and performance was significantly worse than when melodies were in separate ranges—the same exact ranges used in Experiment 1 here. This was likely due

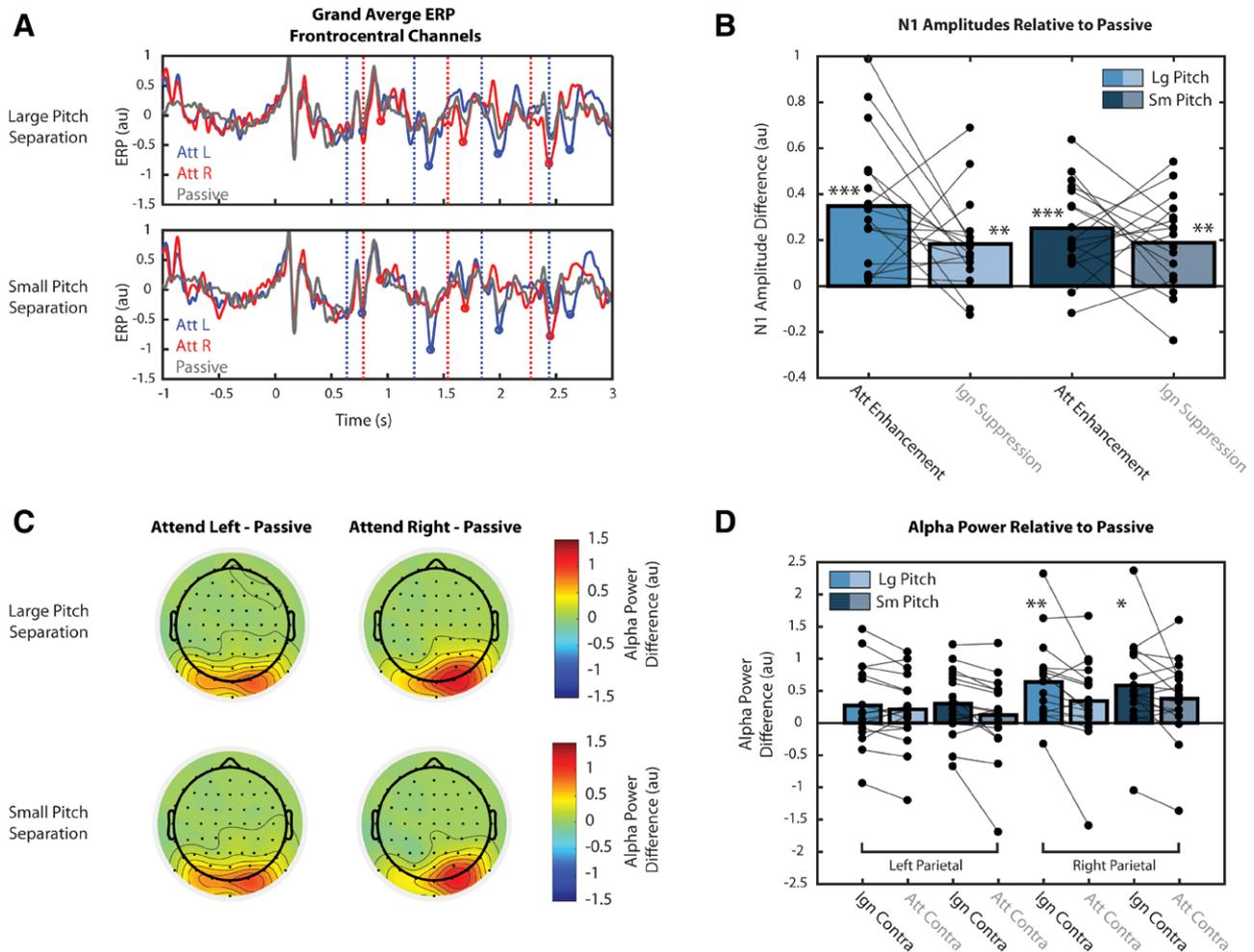


Fig. 5. Event-related potentials (ERPs) and alpha power during passive trials for Experiment 2. Single, double, and triple asterisks indicate quantities significantly greater than zero at the 0.05, 0.01, and 0.001 significance levels, respectively. A, Grand average ERPs from Experiment 2 with passive trial average (gray trace) superimposed over attend-left (blue) and attend-right (red) trials. B, Effects of attention on the magnitude of N1 responses, relative to passive conditions, calculated as Attentional Enhancement and Ignore Suppression. C, Effects of attention on the strength of parietal alpha oscillations, relative to passive conditions, averaged over the cue period. D, Difference in alpha power with respect to the passive condition, calculated separately for left and right parietal sensors.

to difficulty segregating the competing streams. The fact that we did not observe degraded N1 modulation in the small pitch separation condition was likely due to the fact that competing melodies did not have overlapping pitch ranges as in Choi et al. (2014), but distinct ranges that were close together (~1 semitone difference). This design difference, and the fact that behavioral measures show that subjects had performed well on the task in all conditions, suggests that subjects were able to segregate and select targets regardless of the available pitch cues.

The fact that the N1 was modulated similarly does not mean that spatial features were used in the same way to maintain attention across conditions. In fact, there are a number of experiments that show N1 modulation in response to attended auditory stimuli that were not spatially separate from competing objects (Hansen and Hillyard 1988; Kong et al. 2014). Thus, the N1 serves as an index of selective attention independent of the features used for selection. If we wish to index the extent to which spatial features are used to direct top-down attention, then measuring the N1 is insufficient if other features can also be used. Instead, we look to modulation of alpha power, which occurs over cortical regions that map space. If spatial features

are used to a lesser degree to focus attention, then we may expect reduced attentional modulation of parietal alpha power.

Lateralization of Parietal Alpha Power Reflects Spatial Focus of Auditory Selective Attention

While parietal alpha has been studied extensively as a correlate of visuospatial attention, its role in auditory spatial attention is less clear. Nonetheless, growing evidence supports the idea that auditory spatial attention recruits the same cortical networks that are active during visual spatial attention. Early neuroimaging studies defined a dorsal frontoparietal network responsible for orienting visual attention to a particular location (Posner & Petersen 1990; Corbetta & Shulman 2002; Petersen & Posner 2012). This network was composed of the frontal eye fields and superior parietal lobe. Later studies revealed that this network is also involved in auditory attention (Lewis et al. 2000; Shomstein & Yantis 2006; Krumbholz et al. 2009; Braga et al. 2013), but did not establish whether the network was truly supramodal or was instead composed of modality-specific subnetworks.

Recent functional magnetic resonance imaging studies have identified interleaved visual and auditory-biased networks in lateral frontal cortex (Michalka et al. 2015; Noyce et al. 2017), suggesting that there are modality-specific networks for attention. The visual-biased network contains superior and inferior precentral sulcus, which are functionally connected to posterior visual sensory regions; the auditory-biased regions contain transverse gyrus intersecting precentral sulcus and caudal inferior frontal sulcus, which are functionally connected to posterior auditory sensory regions. Although these networks are modality specific, the visual-biased network is flexibly recruited during auditory attention when spatial focus is required to perform the task (Michalka et al. 2015, 2016; Noyce et al. 2017). When the task has high temporal demands, the auditory-biased network is active in both vision and audition. Thus, while there are modality-specific networks for attention, these networks are recruited in a non-modality-specific manner depending on the attended features (spatial versus temporal).

If the same frontoparietal network underlies auditory and visual spatial attention, then one would expect to observe the same EEG correlates of spatial attention over parietal cortex during spatial attention independent of stimulus modality. Therefore, if increased parietal alpha reflects suppression of unattended space in vision (Worden et al. 2000; Sauseng et al. 2005; Kelly et al. 2006; Foxe & Snyder 2011; Händel et al. 2011; Payne et al. 2013), then it should be present for spatial suppression in audition. Indeed, at least one previous study has shown evidence of parietal alpha modulation during auditory spatial attention (Banerjee et al. 2011).

Our results are consistent with these findings. We observed that alpha was lateralized after subjects were given a spatial cue, and this lateralization pattern reflected the space being ignored (i.e., alpha was greater ipsilateral to the attended location). The results from Experiment 1 suggest that subjects at least initially oriented top-down attention using known spatial features of the target even if they could depend solely on pitch information to perform the task. In Experiment 2, subjects *had* to initially orient attention in space due to the absence of pitch cues. Therefore, the observed alpha modulation during the cue period in this experiment strengthens the argument that parietal alpha lateralization reflects the use of *spatial* features to help focus attention. Furthermore, our comparison of parietal alpha oscillations between active and passive conditions during the cue period supports the view that alpha reflects active suppression of contralateral stimuli.

Alpha Lateralization Is Weak When Pitch Cues are Strong, Reflecting the Fact That Pitch Can Also Be Used to Help Focus Attention

While space is first coded at the level of the retina in vision, the auditory system relies on interaural time and level differences to localize sound along the azimuth. Therefore, the mechanisms by which auditory attention operates are likely not inherently spatial. This explains why in the vision literature, spatial and feature-based (e.g., color, texture, etc.) attention are described separately, yet in audition, perceived location is described as a feature itself (Shinn-Cunningham 2008; Maddox & Shinn-Cunningham 2012; Shinn-Cunningham et al. 2017). In audition, nonspatial features, such as pitch, can often be used to direct and maintain attention to an ongoing stream (Lee et

al. 2012, 2014; Maddox & Shinn-Cunningham 2012). If these pitch cues are large compared with the available spatial cues, then individuals may depend more on pitch as the feature on which to base attention. When pitch cues are less informative, however, it may be more beneficial to depend on spatial differences among competing stimuli to maintain attention.

If parietal alpha truly reflects the use of spatial features during sustained top-down attention, then its modulation should be weaker during tasks in which spatial features are redundant with other nonspatial features. Our results support this view. As argued earlier, alpha lateralization occurred during the cue period in all conditions, which suggests that spatial attention was initially directed using the known spatial features. During the stimulus period, however, this lateralization was weak (i.e., not significantly greater than zero) when strong pitch cues were available (i.e., Experiment 1 and the large pitch separation condition of Experiment 2). In Experiment 2, we also observed that this lateralization was significantly larger in the small pitch separation condition than in the large pitch separation conditions. These results likely reflect the fact that, in addition to space, pitch cues could also be used to differentiate target from distractor. Therefore, even though subjects initially directed attention to the location of interest, once the auditory object was selected, its pitch was used to maintain attention throughout the remainder of the stream. When these pitch cues were weak, spatial features may have been necessary to maintain attention, which is why we observed alpha lateralization throughout the small pitch separation trials.

In addition to pitch and space, the distinct timings of notes in each stream could also be used for stream segregation. However, because these timings were kept constant across Experiments 1 and 2, we believe that this factor does not contribute to any of the differences in alpha power we observed across conditions. Furthermore, even though the temporal pattern could be used to attend selectively to a target stream, which theoretically could allow listeners to rely even less on spatial cues, we nonetheless observed alpha lateralization in the small pitch separation condition.

The Degree of Alpha Lateralization Does Not Explain Performance

In Experiment 2, we observed differences in performance between pitch separation conditions. Therefore, it may be possible that the differences in alpha lateralization observed between the two conditions are due to differences in ability to perform the task instead of differences in pitch cue strength. However, we argue that this is not the case for two reasons. First, we removed all trials in which subjects responded incorrectly, so we assume that the EEG signal we observed was recorded when subjects successfully focused attention and not when they may have been struggling to do so. Second, if it were the case that alpha lateralization was stronger because the small pitch separation condition was more difficult, then we may expect some correlation of AMI_{α} with performance measures—subjects who are inherently worse at the task may require more suppression of distractors, which may manifest in greater alpha lateralization. However, we observed no correlation of performance measures with alpha modulation in either left or right parietal channels. Furthermore, that a similar lateralization pattern

was observed during the cue period for both large and small pitch separation conditions despite the performance differences suggests that alpha is indexing spatial focus of attention and not task difficulty.

One may still argue that the greater alpha lateralization observed during the stimulus period was due to more effort being required in the small pitch separation condition, even though performance measures were not correlated with lateralization measures. While this may be true, we argue here that the increased effort required may be defined as the greater need to orient and maintain attention in space because pitch cues are less informative. Therefore, more effort here means more use of spatial attention, which is reflected by stronger alpha lateralization. In the future, more efforts should be made to disentangle the effects of task difficulty and spatial attention on parietal alpha power.

Caveats: Weighting the Effects of Pitch and Space

While our results suggest that the degree of alpha lateralization reflects the degree to which spatial features are used to selectively attend, we did not parametrically adjust spatial and pitch separations of competing melodies. Rather, we tested two conditions in which the pitch separation was different while a somewhat small spatial separation (± 100 μ sec ITD) was held constant. Therefore, in this study, we assumed that when pitch differences are larger, less dependence on spatial features is required during the course of the auditory stream. It may be the case, however, that given the same pitch separation, spatial features would be used to a greater extent if the spatial separation were larger. In fact, previous studies have shown that both pitch and perceived location have similar effects on ability to selectively attend (Maddox & Shinn-Cunningham 2012), with performance improving as the task-relevant feature separation increased. Future work should aim to address under what conditions and to what degree alpha is lateralized given the available space and pitch cues.

CONCLUSIONS

Understanding which features are used during auditory attention advances our understanding of how individuals communicate in noisy environments. In these settings, individuals must not only be able to segregate various auditory objects but also select a single relevant object while suppressing distractors. Our results show that while N1 modulation reflects selection during both spatial and nonspatial auditory attention, parietal alpha lateralization reflects the degree to which spatial features are used to suppress distractors. These results demonstrate the degree to which EEG correlates of spatial attention could be used when designing a noninvasive method for tracking listening focus in a variety of complex auditory scenes—scenes which contain different levels of spatial and nonspatial features. In addition to behavioral studies, such EEG studies could reveal under which conditions spatial features are used to help solve the cocktail party problem (Cherry 1953).

Communication at the cocktail party is particularly challenging for those with hearing loss, even when assistive devices are worn (Marrone et al. 2008). There is a need for technologies that assist object selection in complex scenes. Proposed strategies involve predicting where individuals intend to direct attention to enhance selection of objects at that location (Shinn-Cunningham & Best 2008; Kidd Jr et al. 2013). Such

predictions may be made using measures from noninvasive EEG. However, to predict where individuals intend to focus attention using correlates of spatial attention, we have to know that spatial attention is being used in the first place. Our results suggest that the degree of parietal alpha lateralization may reflect the degree to which spatial features are used during attention, and so if an individual is attempting to orient attention using nonspatial features, then alpha lateralization would not be informative. This technique would instead have to rely on other EEG correlates such as the N1, which require knowledge of the target object's temporal structure. Furthermore, individuals with hearing loss often have degraded object representations beginning at the level of auditory periphery (Shinn-Cunningham & Best 2008; Dai et al. 2018) which may degrade the ability to use spatial features for focusing attention. Future work should explore the degree to which alpha is lateralized in listeners with hearing loss performing a spatial attention task.

In this study, we aimed to determine if lateralization of parietal alpha power reflected the use of spatial features during auditory selective attention. Our results showed that given a spatial cue, alpha was initially lateralized to reflect the location of the to-be-ignored auditory stream. We measured whether this lateralization would persist over the course of an auditory stream if strong pitch cues differentiated target from distractor. Our results showed that when pitch cues were strong, alpha lateralization was weakened after the target began to play, reflecting the fact that pitch could also be used to help focus attention. These results show that even when spatial attention is used initially to focus attention on a target, maintenance of attention can be accomplished using nonspatial cues when other acoustic features differentiate target from distractor streams.

ACKNOWLEDGMENTS

We thank Jasmine Kwasa for helping with data collection for Experiment 1.

This work was supported by National Institutes of Health (NIH) Award No. 1R01DC013825. L.M.B. was supported by NIH Computational Neuroscience Training Grant No. 5T90DA032484-05.

The authors have no conflicts of interest to disclose.

Address for correspondence: Barbara G. Shinn-Cunningham, Neuroscience Institute, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA. E-mail: bgsc@andrew.cmu.edu

Received May 5, 2019; accepted February 27, 2020.

REFERENCES

- 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? *J Neurosci*, *31*, 9923–9932.
- Best, V., Ozmeral, E. J., Kopco, N., Shinn-Cunningham, B. G. (2008). Object continuity enhances selective auditory attention. *Proc Natl Acad Sci U S A*, *105*, 13174–13178.
- Braga, R. M., Wilson, L. R., Sharp, D. J., Wise, R. J., Leech, R. (2013). Separable networks for top-down attention to auditory non-spatial and visuospatial modalities. *Neuroimage*, *74*, 77–86.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spat Vis*, *10*, 433–436.
- Bressler, S., Goldberg, H., Shinn-Cunningham, B. (2017). Sensory coding and cognitive processing of sound in Veterans with blast exposure. *Hear Res*, *349*, 98–110.
- Chaumon, M., Bishop, D. V., Busch, N. A. (2015). A practical guide to the selection of independent components of the electroencephalogram for artifact correction. *J Neurosci Methods*, *250*, 47–63.

- Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *J Acoust Soc Am*, 25:975–979.
- Choi, I., Rajaram, S., Varghese, L. A., Shinn-Cunningham, B. (2013). Quantifying attentional modulation of auditory-evoked cortical responses from single-trial electroencephalography. *Front Hum Neurosci*, 7, 115.
- Choi, I., Wang, L., Bharadwaj, H., Shinn-Cunningham, B. (2014). Individual differences in attentional modulation of cortical responses correlate with selective attention performance. *Hear Res*, 314, 10–19.
- Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci*, 3, 201–215.
- Dai, L., Best, V., Shinn-Cunningham, B. G. (2018). Sensorineural hearing loss degrades behavioral and physiological measures of human spatial selective auditory attention. *Proc Natl Acad Sci U S A*, 115, E3286–E3295.
- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods*, 134, 9–21.
- Deng, Y., Reinhart, R. M., Choi, I., Shinn-Cunningham, B. G. (2019). Causal links between parietal alpha activity and spatial auditory attention. *eLife*, 8, e51184.
- Foxe, J. J., & Snyder, A. C. (2011). The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention. *Front Psychol*, 2, 154.
- Händel, B. F., Haarmeier, T., Jensen, O. (2011). Alpha oscillations correlate with the successful inhibition of unattended stimuli. *J Cogn Neurosci*, 23, 2494–2502.
- Hansen, J. C., & Hillyard, S. A. (1988). Temporal dynamics of human auditory selective attention. *Psychophysiology*, 25, 316–329.
- Jung, T. P., Makeig, S., Humphries, C., Lee, T. W., McKeown, M. J., Iragui, V., Sejnowski, T. J. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, 37, 163–178.
- Kayser, J., & Tenke, C. E. (2006). Principal components analysis of Laplacian waveforms as a generic method for identifying ERP generator patterns: I. Evaluation with auditory oddball tasks. *Clin Neurophysiol*, 117, 348–368.
- Kayser, J., & Tenke, C. E. (2015). On the benefits of using surface Laplacian (current source density) methodology in electrophysiology. *Int J Psychophysiol*, 97, 171–173.
- 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. *J Neurophysiol*, 95, 3844–3851.
- Kidd, G. Jr, Favrot, S., Desloge, J. G., Streeter, T. M., Mason, C. R. (2013). Design and preliminary testing of a visually guided hearing aid. *J Acoust Soc Am*, 133, EL202–EL207.
- Kong, Y. Y., Mullangi, A., Ding, N. (2014). Differential modulation of auditory responses to attended and unattended speech in different listening conditions. *Hear Res*, 316, 73–81.
- Kreitewolf, J., Mathias, S. R., Trapeau, R., Obleser, J., Schönwiesner, M. (2018). Perceptual grouping in the cocktail party: Contributions of voice-feature continuity. *J Acoust Soc Am*, 144, 2178.
- Krumbholz, K., Nobis, E. A., Weatheritt, R. J., Fink, G. R. (2009). Executive control of spatial attention shifts in the auditory compared to the visual modality. *Hum Brain Mapp*, 30, 1457–1469.
- Larson, E., & Lee, A. K. (2014). Switching auditory attention using spatial and non-spatial features recruits different cortical networks. *Neuroimage*, 84, 681–687.
- Lee, A. K., Larson, E., Maddox, R. K., Shinn-Cunningham, B. G. (2014). Using neuroimaging to understand the cortical mechanisms of auditory selective attention. *Hear Res*, 307, 111–120.
- Lee, A. K., Rajaram, S., Xia, J., Bharadwaj, H., Larson, E., Hämäläinen, M. S., Shinn-Cunningham, B. G. (2012). Auditory selective attention reveals preparatory activity in different cortical regions for selection based on source location and source pitch. *Front Neurosci*, 6, 190.
- Lewis, J. W., Beauchamp, M. S., DeYoe, E. A. (2000). A comparison of visual and auditory motion processing in human cerebral cortex. *Cereb Cortex*, 10, 873–888.
- Maddox, R. K., & Shinn-Cunningham, B. G. (2012). Influence of task-relevant and task-irrelevant feature continuity on selective auditory attention. *J Assoc Res Otolaryngol*, 13, 119–129.
- Marrone, N., Mason, C. R., Kidd, G. Jr. (2008). Evaluating the benefit of hearing aids in solving the cocktail party problem. *Trends Amplif*, 12, 300–315.
- McFarland, D. J. (2015). The advantages of the surface laplacian in brain-computer interface research. *Int J Psychophysiol*, 97, 271–276.
- Mehraei, G., Shinn-Cunningham, B., Dau, T. (2018). Influence of talker discontinuity on cortical dynamics of auditory spatial attention. *Neuroimage*, 179, 548–556.
- Michalka, S. W., Kong, L., Rosen, M. L., Shinn-Cunningham, B. G., Somers, D. C. (2015). Short-term memory for space and time flexibly recruit complementary sensory-biased frontal lobe attention networks. *Neuron*, 87, 882–892.
- Michalka, S. W., Rosen, M. L., Kong, L., Shinn-Cunningham, B. G., Somers, D. C. (2016). Auditory spatial coding flexibly recruits anterior, but not posterior, visuotopic parietal cortex. *Cereb Cortex*, 26, 1302–1308.
- Noyce, A. L., Cestero, N., Michalka, S. W., Shinn-Cunningham, B. G., Somers, D. (2017). Sensory-biased and multiple-demand processing in human lateral frontal cortex. *J Neurosci*, 37, 8755–8766.
- O’Sullivan, J., Chen, Z., Herrero, J., McKhann, G. M., Sheth, S. A., Mehta, A. D., Mesgarani, N. (2017). Neural decoding of attentional selection in multi-speaker environments without access to clean sources. *J Neural Eng*, 14, 056001.
- Payne, L., Guillory, S., Sekuler, R. (2013). Attention-modulated alpha-band oscillations protect against intrusion of irrelevant information. *J Cogn Neurosci*, 25, 1463–1476.
- Petersen, S. E., & Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annu Rev Neurosci*, 35, 73–89.
- Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annu Rev Neurosci*, 13, 25–42.
- 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. *Eur J Neurosci*, 22, 2917–2926.
- Shinn-Cunningham, B. G. (2008). Object-based auditory and visual attention. *Trends Cogn Sci*, 12, 182–186.
- Shinn-Cunningham, B. G., & Best, V. (2008). Selective attention in normal and impaired hearing. *Trends Amplif*, 12, 283–299.
- Shinn-Cunningham, B., Best, V., Lee, A. K. (2017). Auditory object formation and selection. In J. C. Middlebrooks & J. Z. Simon (Eds.), *The Auditory System at the Cocktail Party* (pp. 7–40). Springer.
- Shomstein, S., & Yantis, S. (2006). Parietal cortex mediates voluntary control of spatial and nonspatial auditory attention. *J Neurosci*, 26, 435–439.
- Tune, S., Wöstmann, M., Obleser, J. (2018). Probing the limits of alpha power lateralisation as a neural marker of selective attention in middle-aged and older listeners. *Eur J Neurosci*, 48, 2537–2550.
- Van Eyndhoven, S., Francart, T., Bertrand, A. (2017). EEG-informed attended speaker extraction from recorded speech mixtures with application in neuro-steered hearing prostheses. *IEEE Trans Biomed Eng*, 64, 1045–1056.
- van Noorden, L. P. A. S. (1975). *Temporal Coherence in the Perception of Tone Sequences* [PhD thesis]. Eindhoven University of Technology.
- Worden, M. S., Foxe, J. J., Wang, N., Simpson, G. V. (2000). Anticipatory biasing of visuospatial attention indexed by retinotopically specific-band electroencephalography increases over occipital cortex. *Journal of Neuroscience*, 20, 1–6.
- Wöstmann, M., Herrmann, B., Maess, B., Obleser, J. (2016). Spatiotemporal dynamics of auditory attention synchronize with speech. *Proc Natl Acad Sci U S A*, 113, 3873–3878.