Brain areas associated with visual spatial attention display topographic organization during auditory spatial attention

Tzvetan Popov[1,2], Bart Gips[3], Nathan Weisz[4], Ole Jensen[5]

Author's affiliations:

1. Methods of Plasticity Research, Department of Psychology, University of Zurich, Zurich, Switzerland
2. Department of Psychology, University of Konstanz, Konstanz, Germany
3. NATO Science and Technology Organization Centre for Maritime Research and Experimentation (CMRE) La Spezia, Italy
4. Centre for Cognitive Neuroscience and Department of Psychology, University of Salzburg, Salzburg, Austria.
5. School of Psychology, University of Birmingham, Birmingham, UK

Corresponding author: Tzvetan Popov
e-mail: tzvetan.popov@uzh.ch

Corresponding address: Methods of Plasticity Research Laboratory, Department of Psychology, University of Zurich, Andreasstrasse 15, CH-8050, Zurich, Switzerland

Keywords: auditory, attention, alpha, navigation, spatial orientation, EEG, oscillations

Abstract

It is well-established that power modulations of alpha oscillations (8-14 Hz) reflect the retinotopic organization of visuospatial attention. To what extend this organization generalizes to other sensory modalities is a topic of ongoing scientific debate. Here, we designed an auditory paradigm eliminating any visual input in which participants were required to attend to upcoming sounds from one of 24 loudspeakers arranged in a horizontal circular array around the head. Maintaining the location of an auditory cue was associated with a topographically modulated distribution of posterior alpha power resembling the findings known from visual attention. Alpha power modulations in all electrodes allowed us to predict the sound location in the horizontal plane using a forward encoding model. Importantly, this prediction was still possible, albeit weaker, when derived from the horizontal electrooculogram capturing saccadic behavior. We conclude that attending to an auditory target engages oculomotor and visual cortical areas in a topographic manner akin to the retinotopic organization associated with visual attention suggesting that the spatial distribution of alpha power reflects the supramodal organization of egocentric space.

Introduction

Adaptive behavior in complex environments requires a mechanism enabling the conversion of external events into internal egocentric representations in a goal-directed manner. This includes processes to prioritize goal-relevant stimulus features. In the visual domain, alpha oscillatory activity (8-14 Hz) has been proposed to reflect such a prioritization mechanism. For example, when attention is spatially oriented to a particular external event, alpha power is reduced contralateral to the attended event in a topographic, i.e. retinotopically organized fashion [e.g.[1-3]]. To date, the evidence for such an organization is based on experimental designs using 2D visual displays investigating the effects of elevation and azimuth in the visual field. However, navigation and orientation of organisms within the surrounding environment unfolds within the *full* azimuth plane and relies on multimodal sensory processing particularly integrating the visual and auditory domain. Thus, supramodal spatial mapping is needed to align visual and auditory space to support operations in real-life environments. Here we test the preregistered hypothesis that alpha power modulations in the parietal cortex constitute an overarching supramodal circuit in which a given events' position is represented in an egocentric reference frame common for both vision and audition.

Spatial analysis and discrimination of auditory input is essential for survival for many living organisms and is central to human spatial orientation and social communication in particular. In an audition, interaural loudness and latency differences are utilized to infer the origin of the auditory source in the horizontal plane, with the superior olive transmitting information bilaterally to upstream regions. Where in the brain is visual and auditory spatial information integrated? The parietal cortex has been established as a region encoding the azimuth of auditory cues[4-6]. It is conceivable that incoming auditory input converges on a supramodal egocentric representation of space to be integrated with other information and be made

accessible to action. In audio-visual spatial cueing paradigms both auditory and parietal areas display parietal modulation of alpha activity reminiscent of the ones observed in visual attention paradigms [7-13]. While these modulations in alpha power usually are interpreted as reflecting functional inhibition (power increase) and disinhibition (power decrease) of the visual processing stream it remains unclear to what extend alpha power fluctuations during auditory tasks also reflect the encoding of physical space.

Modulations of alpha activity in relation to space have mainly been investigated within the visual domain. Consequently, numerous studies expand the functional relevance to e.g. working memory, distractor suppression, and episodic memory formation[9,14,15]. The close anatomical proximity of visual and parietal areas contributes to the assumption that modulations of alpha power are primarily linked to vision. To date, studies examining the effects of spatially distributed auditory events independently of visual cues are few. The existing literature is limited to designs where the attention of lateralized auditory material is typically signaled by visual cues or followed by behavioral responses in accordance with some information presented on a visual display. Moreover, the influence of micro-saccades on neural activity has raised some concerns allowing an alternative view on alpha power modulation as putatively explained by micro-saccades to the attended visual field [16-19]. It is conceivable that both the encoding of the visual cue and the preparation to respond to a visual target are associated with micro-saccade generation in the register of the cued location. This, in turn, could give rise to the modulation of parietal alpha activity during the maintenance of auditory spatial information. Is alpha power modulation “supramodal” reflecting the orientation to spatial events in the environment independent of the sensory modality? Or alternatively, even in the context of a purely auditory task, orientation to spatial cues is associated with a consistent visuomotor activity, in turn initiating patterns of alpha power modulations in audition that resemble the once observed during visual-spatial attention?

In the present report, we designed a study involving 24 loudspeakers horizontally positioned around the participant's head. An auditory cueing paradigm was used while high-density EEG was used to acquire the participant's brain activity. The preregistered hypotheses (https://osf.io/kp95j) were:

H1: There is a spatio-temporal pattern of neural activity in the EEG data that will allow decoding the direction of auditory attention.

In support of this hypothesis, we expect that alpha power modulation is independent of the sensory domain: the direction of attention cued by auditory stimuli to the left hand side should prompt modulation of contralateral alpha power over posterior electrodes and vice versa.

H2: Spatial information is encoded following the presentation of auditory cues.

In support of this hypothesis, we reasoned that the decoding performance can be compared between periods of spatial auditory cue maintenance and pre-cue baseline. Going beyond the left-right stimulus presentation, all additional loudspeaker angles will be considered.

To address the contribution of oculomotor activity, exploratory analyses were conducted utilizing the horizontal electrooculogram (hEOG) during the maintenance interval of an auditory spatial cue.

Material and Methods

Participants

Thirty-one undergraduates were recruited at the local university (mean age M±SD 23.6±3.57 years, 18 female). All but one reported no history of neurological and/or psychiatric disorders. All participants gave written informed consent in accordance with the Declaration of Helsinki prior to participation. The study was approved by the University of Konstanz ethics committee.

Stimulus material and procedure

An auditory version of the spatial Eriksen flanker task was used [2,20]. Participants were instructed to maintain a comfortable sitting position in the center of an aluminum ring (Figure 1B). After a baseline period (2 s, Figure 1A), an auditory cue (100 ms duration; 440 Hz) was presented randomly at one of 8 locations (0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°, 360°, Figure 1A). Given an average ear-to-ear distance of approximately 20 cm, the half wavelength of sound waves below 800 Hz is larger than the head size such that phase delays between both ears can be reliably identified. After a delay interval of 2.5 ± 1 s, during which subjects maintained the cued position, a target syllable (German, "goeb" or "goed") appeared at that location, embedded in a circular array of 24 speakers mounted in 15° distance on the inner surface of the aluminum ring. Participants indicated via button press whether the target syllable was a "goeb" (left index finger) or a "goep" (right index finger). All responses were given with the right hand. 160 trials were presented in each of three blocks separated by a short break. The second and third blocks were identical to the first one. The only difference was that the location of cues and targets were shifted with 15° (2nd block) and 30° (3rd block) thereby ensuring a full 24 location circular coverage. Participants were not aware of this

change in speaker arrangement. A total of 480 trials (20 per location) were presented. Stimulus presentation was controlled using Presentation software (www.neurobs.com) on a Windows 7 PC.

Data acquisition

The EEG was measured in an electrically shielded room using a high-density 256-channel EGI system with a HydroCel Geodesic Sensor Net (GSN; Electrical Geodesics, Inc., Eugene, Oregon, USA). Prior to sampling at 1000 Hz, the EEG was filtered using a 0.1 Hz high-pass and a 400 Hz low-pass hardware filters. The vertex (Cz) electrode served as a recording reference. All subsequent analyses were performed after converting the data to a common reference. Following EGI acquisition guidelines, electrode impedances were kept below 30kΩ, which is adequate because of the high input impedance of the EGI amplifiers. Standard positions for the present montage were registered to later align with a Montreal Neurological Institute (ICBM 2009a Nonlinear Asymmetric 1×1×1 mm) template brain (Montreal Neurological Institute, Montreal, Canada http://www.bic.mni.mcgill.ca/ServicesAtlases/ICBM152NLin2009).

Neural data analysis

Data analysis was performed using the MATLAB FieldTrip toolbox[21]. After demeaning and removing the linear trend across the session, an independent component analysis (ICA)[22] was used to remove variance associated with eye blink and cardiac activity. On average 19.4 trials per location (STD = 0.2) were retained for further analyses.

Spectral analysis

Spectral analysis was computed for each trial using a Fast Fourier Transformation (FFT) based on a sliding window of 500 ms multiplied with a Hanning taper resulting in

frequency smoothing of ~3 Hz. Power estimates were calculated for the latency from -1 to 2 s after cue onset in steps of 50 ms. Subsequently, the power estimates were averaged over trials. Analysis of alpha power lateralization was performed based on the trials with left and right most cueing locations (Figure 1A, left speakers 6,7,8 and right speakers 18,19,20).

Source analysis

Source estimates were computed in the time as well as in the frequency domain. In the frequency domain, an adaptive spatial filtering algorithm was used (dynamic imaging of coherent sources, DICS)[23]. This algorithm uses the cross-spectral density matrix from the EEG data and the lead-field derived from the forward model to construct a spatial filter for a specific location. This matrix was calculated using a multi-taper FFT approach for data in the 0.3 – 0.8 s interval following the cue onset. Spectral smoothing of ±2Hz around (Δf 4Hz) center frequencies of 10 Hz was applied including the power in the 8 – 12 Hz (alpha) range. These spectral density matrices and thus the spatial filters were participant-specific and estimated based on all trials and used to estimate the power for the trials with the leftmost (90 ± 15°) and rightmost cues (270 ± 15 °). This so-called common spatial filter based on all trials ensures that potential differences in oscillatory power are not due to differences in filter estimates of conditions. A standard forward model was constructed from the MNI ICBM 2009 template brain (http://www.bic.mni.mcgill.ca/ServicesAtlases/ICBM152NLin2009) using the OpenMEEG[24] implementation of the Boundary Element Method (BEM). A parcellation scheme based on the Desikan-Kiliani atlas was implemented[25]. A cortical surface source model was generated consisting of 2002 dipole locations. The forward solution was applied to all participants.

In the time domain, a related spatial filtering algorithm (LCMV, linearly constrained minimum variance) was used[26]. This algorithm uses the covariance matrix of the EEG data to

construct a spatial filter for a given location. The covariance matrices for these spatial filters were estimated based on all trials and -0.3 – 1 s interval with respect to cue onset. A 1 – 20 Hz bandpass filter was applied before these operations. These filters were applied to the scalp data to derive the time series for a given location.

Forward encoding modeling

Forward encoding modeling followed the procedure described in[14] and publicly shared on the https://osf.io/vw4uc/ platform. Briefly, the general assumption is that oscillatory power quantified at each electrode reflects the weighted sum of twenty-four hypothetical directions reflecting the macroscopic manifestation of spatially tuned neuronal populations. Each of these neuronal ensembles is tuned to a different speaker direction (Figure 1). The EEG data were partitioned into two blocks (train and test) with similar trial numbers. A ten-fold random generation of multiple block assignments (e.g. test or train) was utilized and the outcome was averaged over folds. Single-trial alpha power was estimated using a Hilbert transform on the bandpass filtered data (8 -12Hz) identical to the procedures described in (Foster et al., 2016; Foster et al., 2017). To infer the position of the maintained spatial location from the EEG data, a set of 24 basis functions coding for 24 equally spaced directions between 0 and 360° was constructed first. For each time point training data *B1* allowed the estimation of weights that approximated the relative contribution of the 24 hypothesized spatial channels (*k*) to the measured scalp data. Let *B1* (*m* electrodes × *n1* trials) be the signal at each electrode and trial in the training set, *C1* (*k* spatial channels × *n1* trials) the predicted response of each spatial channel, and *W* (*m* electrodes × *k* spatial channels) the weight matrix allowing the linear mapping from "spatial channels space" to electrode space. This mapping was based on a linear model of the form:

$$C_1 = WB_1 + \epsilon \qquad (1)$$

Where ϵ contains (assumed Gaussian) error terms that should be minimized. To this end, ordinary least-squares regression was used to estimate the weight matrix $W\ (m \times k)$:

$$\widehat{W} = C_1 B_1^T (B_1 B_1^T)^{-1} \quad (2)$$

On the basis of this weight matrix and on the test data B_2 (m electrodes × n_2 trials) an estimated response matrix C_2 (k spatial channels × n_2 trials) was calculated:

$$\hat{C}_2 = \widehat{W} B_2 \quad (3)$$

The estimated responses were circularly shifted such that estimates associated with directions that evoked a response were positioned at 0° of the direction space spanning -180° to 180°. Following this step an accurate model is characterized by a maximum at 0° and a minimum at -180°/180° (Figure 4A). In contrast, an inaccurate model fit approximates a flat line. This procedure was performed for each sample point in -1 – 1 s interval with respect to the cue onset. This was repeated until each block had served as a training and test set.

Finally, to interpret the weight matrix W in terms of source origin, an activation matrix A of a corresponding forward encoding model was computed[27]:

$$A = \Sigma_{B_1} W^T \Sigma_{\hat{C}_1} \quad (4)$$

Here, $\Sigma_{B_1} = cov(B_1)$ and $\Sigma_{\widehat{C_1}} = cov(\hat{C}_1) = cov(\widehat{W} B_1)$ are covariance matrices. The advantage of using A instead of the raw weights W is that elements of W may reflect suppression of "signals of no interest"[27]. For example, correlations across sensors in B_1 could be confounded by noise. Therefore, they do not reflect brain activity related to C_1. Transforming to activation patterns A mitigates this problem. Forward encoding modeling was performed in source space in order to map the activation patterns onto the individual brain volume.

Inferential statistical analysis

Quantification of oscillatory measures for inferential statistics was carried out by a cluster-based approach based on randomization [28].This approach identifies clusters (in time, frequency and space, i.e., electrodes) of activity on the basis of whether the null hypothesis can be rejected while addressing the multiple-comparison problem. Relationships between

behavioral (RT) and neural data (tuning response) were examined using Spearman's rank-order correlations (Rho). Rain cloud plots[29] were utilized for data visualization when appropriate.

Results

During EEG acquisition participants were cued to a particular speaker location. After a delay interval, during which maintenance of the cued location was required, a target has been presented at the cued speaker. Participants were asked to indicate via button press whether they heard the syllable "goeb" (left button press) or "goep" (right button press) (Figure 1 A, B). Response times (RT) did not vary with speaker location (Figure 1C) and the overall hit rate was 96.3%±8.3% (M±STD). Behavioral results confirming participant's task compliance and indicate no behavioral bias towards any particular speaker location.

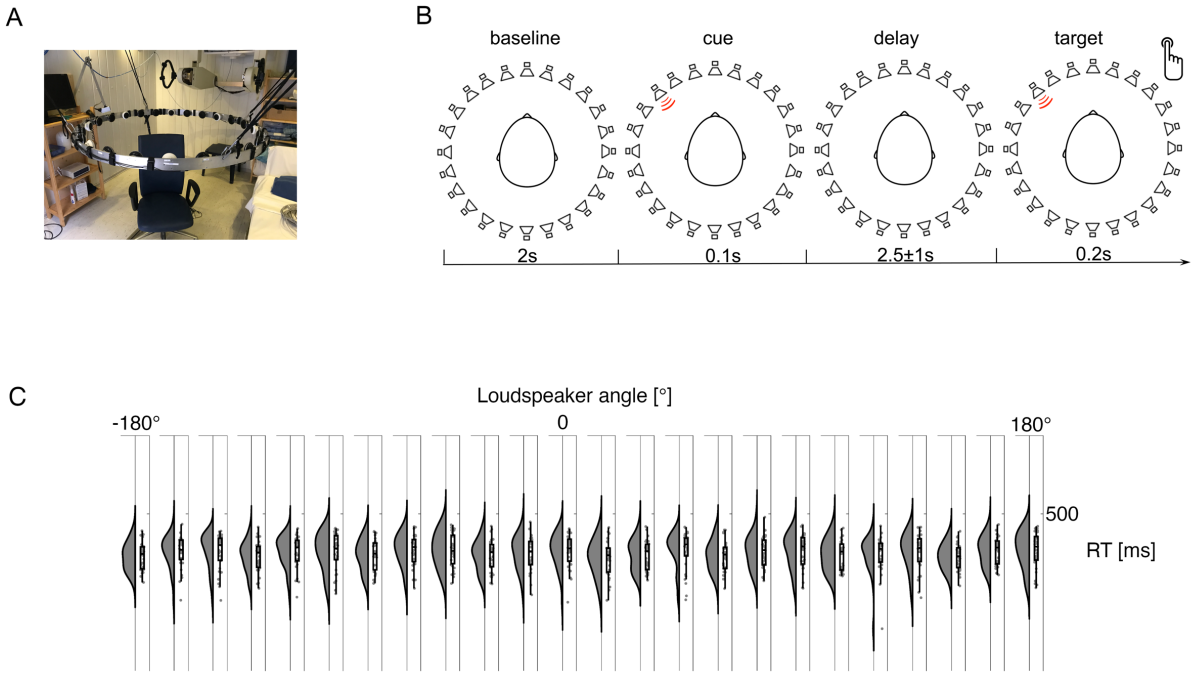


Figure 1: Experimental setup and behavioral results. A – A photograph of the hardware illustrating an aluminum ring holding the 24 loudspeakers equidistantly placed with an angular distance of 15°. Participants were sitting in the chair with the head positioned in the center of the ring. **B-** After a baseline interval of 2 s an auditory cue is presented at one of 24 speaker directions for 100 ms. During the delay interval of 2.5+/-1 s participant's maintained the cued direction in memory. After this delay interval, a target syllable "goeb" or "goep" was presented for 200 ms at the cued direction. Participants were asked to indicate via button press, as fast as possible, whether they heard "goeb" or "goep". **C-** Rain cloud plots per loudspeaker direction illustrating a similar distribution of RT across participants.

The auditory cue presentation was associated with reliable event-related potentials (ERPs) with a typical auditory scalp topography characterized by the largest negativity of the N100 ERP components around the vertex electrode (Figure 2 A). Source analysis revealed the difference in neural generators in the interval 110 –180 ms associated with left vs. right spatial cue processing predominantly in bilateral higher-order auditory and parietal brain areas (Figure 2B). Processing of left auditory cues was associated with a stronger neuronal response in the right parietal cortex contralateral to the cued direction and vice versa.

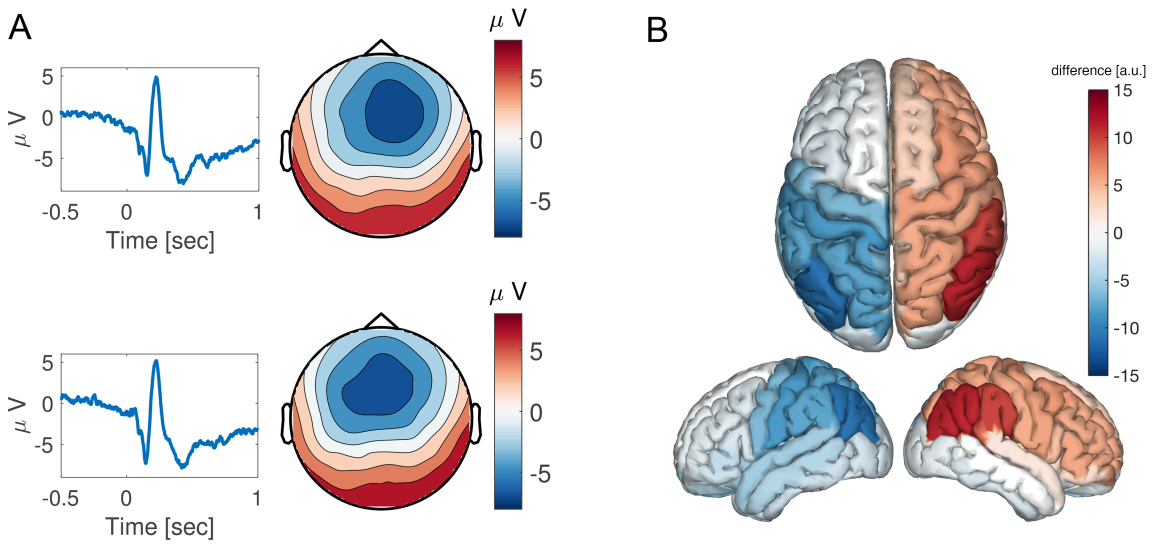


Figure 2: Spatial auditory cues engage the posterior parietal cortex as measured by auditory evoked potentials. A- Time-series illustrate the auditory potentials averaged across participants at electrode Cz, for the left presented stimuli (top) and right presented stimuli (bottom). Cue onset is denoted at 0 sec. The topographies correspond to the latency of the N1 evoked response ~100 ms post cue onset. Cold colors reflect negative and warm colors positive voltage with respect to the pre-cue baseline. **B-** Source reconstruction of the difference left minus right presented auditory stimuli for the time interval 110-180ms post cue onset.

The lateralization in neuronal activity was also apparent when analyzing the data in the time-frequency domain (Figure 3). Maintenance of auditory cues to the left was associated with a contralateral decrease in alpha power (Figure 3A, $p < 0.025$, cluster permutation test) and a relative increase in the ipsilateral hemisphere. Source analysis confirmed lateralized activation of parietal brain areas (Figure 3B), largely resembling the distributed activity observed in the time domain source analysis (Figure 2B). Figure 3C illustrates the time course

of ERP and alpha power lateralization computed from parietal electrodes. Descriptively, the time course of the ERP lateralization closely matches the time course of alpha power lateralization with a reliable minimum in the latency 300-500 ms post cue onset. This overlap is in line with the notion that slow evoked components manifest as a consequence of asymmetric modulation of peaks and troughs of ongoing alpha activity [30-32]. Yet, some reports suggest that event-related responses and alpha activity are not always related[33]. Therefore we calculated the correlation between the lateralization of evoked activity (ERP LI lateralization index, i.e. [left - right]/[left + right]) and alpha power modulation (ALI alpha lateralization index, i.e. [left - right]/[left + right]) across participants (Figure 3C). For each participant and within the window of 300- 500 ms post cue onset, the most negative value of the ERP LI and ALI time courses were identified. Participants exhibiting strong ERP lateralization were also characterized by strong alpha power lateralization (Figure 3D; $\rho = .48$, $p < 0.0058$). In summary, both time and time-frequency domain analyses confirmed the hypothesis that the modulation of neuronal activity in the posterior parietal cortex reflects the maintenance of auditory spatial information. However, it is thus far unclear whether this alpha modulation is associated with only a coarse left versus right differentiation or whether the engagement by auditory attention exploits the spatial high fidelity of the posterior parietal cortex. For this purpose, we aimed to decode the speaker location on the basis of the alpha oscillatory activity.

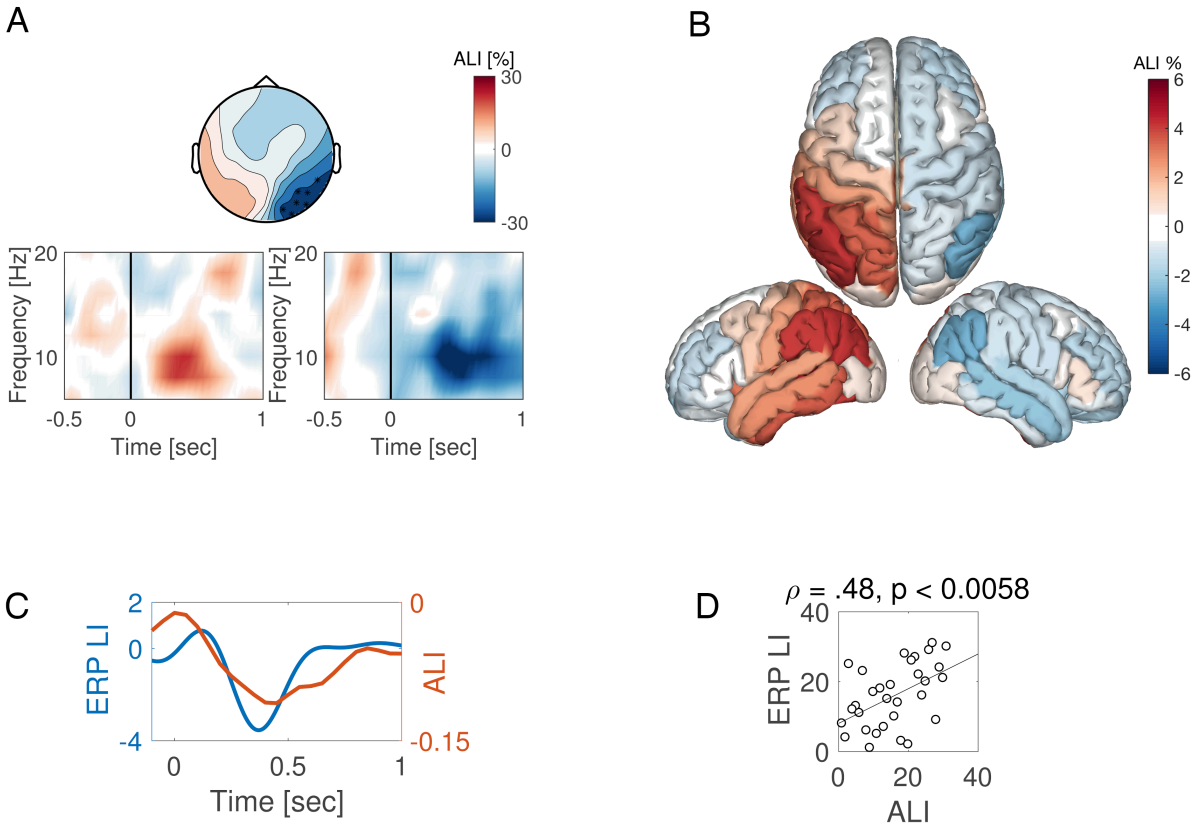


Figure 3: Alpha activity in posterior parietal cortex is modulated during the maintenance of the spatial position of auditory cues. A- Time-frequency representation of power illustrating the contrast (Alpha Lateralization Index; ALI- *[(left-right)/(left+right)]*100*) between left and right presented stimuli. Time is depicted on the x-axis and frequency on the y-axis. The variation of ALI in % is color-coded. Marked electrodes in the topography correspond to the cluster of electrodes confirming significant condition differences in alpha power (cluster permutation approach, $p < 0.025$). **B-** Source reconstruction of the contrast in A illustrating the involvement of posterior-parietal brain areas. **C-** Alpha lateralization index (ALI, orange) and ERP lateralization index (ERP LI) as a function of time. The strongest modulation is apparent ~ 400 ms post cue onset at 0 sec. **D-** The scatterplot illustrates the Spearman correlation between ERP LI and ALI across participants.

Next, we used a forward encoding modeling approach (see Method section) to decode the direction of the cue from the multivariate data in the alpha band (Figure 4). Throughout the delay interval, a robust tuning response to loudspeaker location was observed with a peak latency between ~300-800 ms after cue onset (Figure 4A). This tuning was specific to the delay period as confirmed by a cluster permutation test when comparing to a pre-cue baseline of equal length (i.e. 1000 ms, Figure 4B, cluster permutation test, $p < 0.025$). Tuning response data from the first cluster in Figure 4B (e.g. 435-470 ms and -4° - 14°) was extracted and related to reaction time utilizing Spearman's rank-order correlation (Figure 4C). Participants with strong tuning to speaker location during the delay interval were faster in responding to

the target several seconds later. In summary, analyzing power modulations of alpha activity can reliably decode the loudspeaker location towards which individuals attend, beyond the left-right locations.

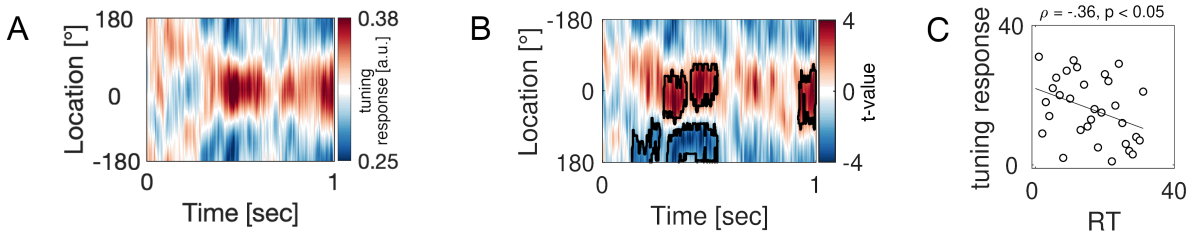


Figure 4: Alpha activity induced by auditory cues reflects the neural encoding of space. A- The tuning response as a function of time-averaged across participants. The X-axis denotes time with cue onset at 0 sec. and y-axis illustrates spatial location ranging from -180° to 180° (see methods). Maximum tuning response (reflected by warm colors) at location 0° corresponds to a strong link between alpha activity and the encoding of spatial information. **B-** Same as in A but expressed in units of t-values derived from the contrast against the pre-cue onset baseline of equal length (cluster permutation test, $p < 0.025$). The black contour line highlights the time ×location cluster supporting the rejection of the null hypothesis (neural tuning data during baseline and delay intervals do not differ). **C-** Scatter plot illustrating the relationship (Spearman rank correlation) between reaction time (RT, abscissa) and the tuning response (ordinate, data extracted from a positive cluster in C).

Finally, mapping activation patterns ("*A"*; see method section) onto the cortical surface revealed that the tuning response was mainly driven by activity originating from the visual and parietal cortex. Despite a clear auditory task demanding encoding, maintenance, and

processing without relying on visual material, brain areas previously associated with the processing of visual information display “retinotopic” organization during audition.

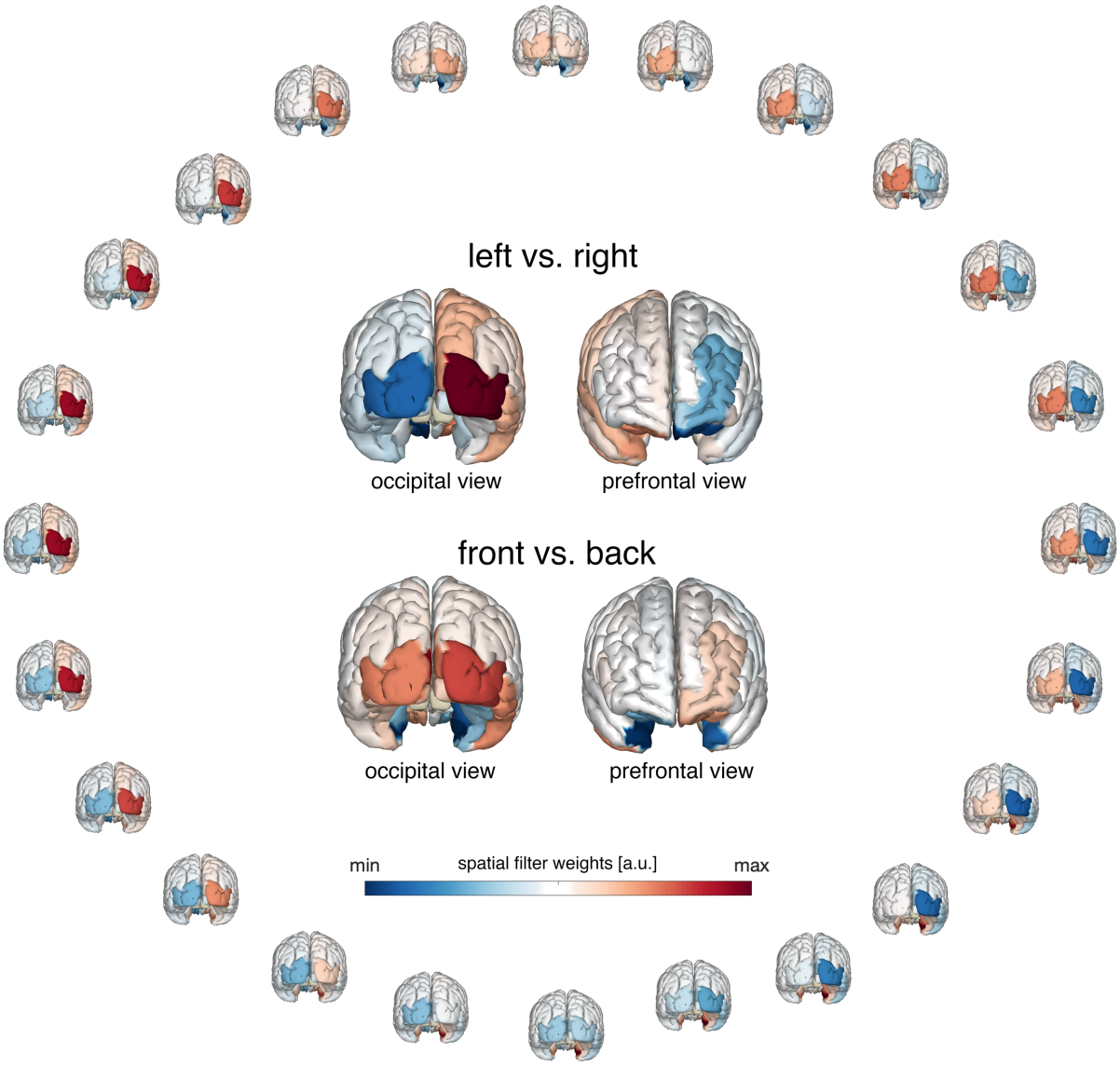


Figure 5: Retinotopic activation of alpha activity in parieto-occipital cortex supports encoding of spatial auditory cues. Distribution of the sources of the alpha power on the cortical surface reflecting the attended direction derived from spatial filter weights obtained from the forward encoding model (activation patterns *A*, see methods). Warm colors indicate a relative alpha band increase and cold colors a decrease expressed in arbitrary units. Insets illustrate the difference between left minus right loudspeaker location (top) and front minus back loudspeaker direction (bottom).

The most informative brain regions are clearly visual. The question arises, why visual cortex activity will contribute to task engagement and processing during the auditory task? As vision is not a required sensory modality a possible interpretation can be derived only from a

multimodal perspective. While a sensory approach would argue for a direct effect of auditory processing on posterior regions (e.g. [34]), recent literature suggests that action-related sensory input mediate multisensory effects. For example, eye movements during auditory attention inform individual group differences within the dorsal attention network[35] and eye-movement related eardrum oscillations link sounds and images in space[36,37]. Thus, alternatively, an affirmative case for the presence of saccades in register of auditory cue location might offer some explanation. We conducted an exploratory analysis re-evaluating the epoched data prior to ICA correction. As an eye-tracking device was not available, we reasoned that if aspects of oculomotor activity are present during the delay interval, these will be reflected in the EEG topography. Specifically, if the saccade direction is consistent towards the direction of the cued position, the difference in ERP topography (left-right) should be characterized by a prototypical saccade topography. The results of this analysis are illustrated in Figure 6 A. The topographic difference in the interval 300 to 1000 ms post auditory cue onset between attention directed towards left speaker location (red) vs. right speaker location (blue) displays a clear oculomotor topography. The ERP time courses a derived from a representative left frontal electrode ('E48') and right frontal electrode ('E221') respectively. The position of these electrodes corresponds to the approximate position of a horizontal electrooculogram (hEOG). Using these two electrodes only, we performed the forward encoding procedure described above. Indeed, an increase in tuning response towards different speaker locations was evident when computed on the basis of the hEOG only (Figure 6B). That is, the variation of saccades during the maintenance interval of the auditory cue was not random but in a direction consistent with the cued speaker location. The effect size of this tuning response was in the range up to a Cohen's d = .8 for a t-value of 2.5 ()38, as compared to the one obtained on the basis of alpha activity (d = 1.44). The alpha tuning response was extracted from the positive cluster illustrated in Figure 4B (location -20 to 15° and latency 285 to 535 ms). Averaging across location degrees and latencies, a single value per participant was derived,

quantifying the strength of the tuning response computed on the basis of alpha activity. These values were related to the tuning response derived from the hEOG electrodes by means of Spearman correlations using a cluster-based permutation approach ($p < 0.05$). Figure 6C illustrates that participants associated with strong alpha tuning were also characterized by strong hEOG-based tuning. That is, saccade and alpha tuning responses were positively related (Figure 6C). This suggests that auditory attention is linked to the visual system, at least in part, thru pro-active orientation towards the relevant sound origin via saccades in the direction consistent with the sound origin.

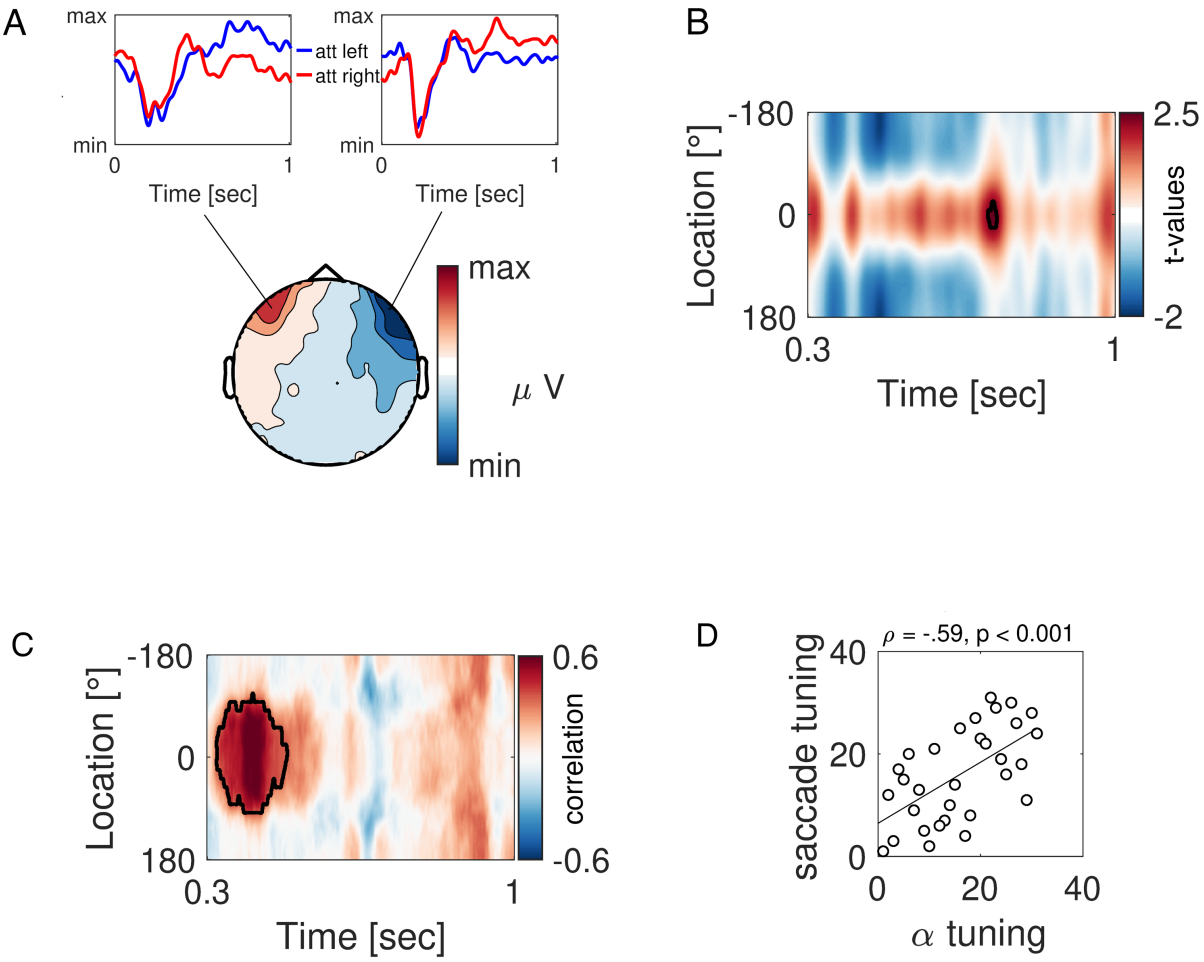


Figure 6: Saccades are consistent with the location maintenance of auditory cues. A: The topography illustrates the ERP difference during the delay interval (300 – 1000 ms, avoiding the first 300 ms dominated by the auditory ERP) between left

vs. right auditory cue location. The time courses reflect the grand average ERP from a representative left ('E48') and right ('E221') electrode. The red time course denotes attention directed towards the right loudspeaker locations and the blue time course towards the left. **B:** The tuning response as a function of time-averaged across participants utilizing only the two electrodes dominated by saccade activity. The X-axis denotes time with cue onset at 0 sec. and y-axis illustrates spatial location ranging from -180° to 180°. Maximum tuning response (reflected by warm colors) at location 0° corresponds to a strong link between saccade direction and the encoding of spatial information. Tuning response is expressed in units of t-values derived from the contrast against the pre-cue onset baseline of equal length (uncorrected, $p < 0.025$). The black contour line highlights the time ×location cluster supporting the rejection of the null hypothesis (neural tuning data during baseline and delay intervals do not differ). **C-** Time × location correlation coefficients between the alpha tuning response (see Figure 4) and the tuning response based on hEOG data (saccade tuning) confirming the relationship that stronger tuning computed on the basis of alpha activity is related to stronger tuning estimated from the hEOG signal. Warm colors denote positive and cold colors negative Spearman's rho correlation coefficients. The black contour line highlights the time ×location cluster supporting the rejection of the null hypothesis (alpha tuning data does not relate to saccade tuning data). **D:** Scatter plot illustrating the relationship highlighted in C between alpha tuning (abscissa) and saccade tuning response (ordinate).

Discussion

Navigation in a complex environment requires the integration of multiple sensory streams. Research has discovered a variety of supramodal brain areas responding to input from different sensory modalities. The sense for direction and with it the deployment of attention relative to the observer's position is egocentric and self-referential. Sense of direction is independent of the sensory modality and can be inferred on basis of vestibular and different sensory input (e.g. visual and/or auditory). The present report provides empirical support for a supramodal neural circuit in service of spatial attention reflected by the spatial distribution of alpha-band activity in the parieto-occipital cortex. In absence of any prior visual input (e.g. typically a visual cue signals spatial auditory information) we observed evoked activity and ongoing alpha power modulation in parieto-occipital brain areas. The activation patterns resembled those known from visual-spatial attention studies and demonstrate a supramodal topographic organization with respect to the direction of attention, initiated at least in part through ocular-motor action. Based on these patterns of neuronal activity in the alpha band, we demonstrate that the maintained spatial direction of the cue can be decoded. In other words, this spatial tuning reflects the aforementioned egocentric, self-referenced nature of spatial attention and orientation in service of overt behavior. In support of this notion, we further observed that stronger spatial tuning was associated with faster responses.

Alpha power modulation in audition encipher physical space

In previous studies, parieto-occipital alpha power modulation in audition has mainly been interpreted as an indicator of active functional inhibition of potentially interfering input from the visual processing stream[38,39] . This is due to the fact that these studies typically involve participant's interaction with the material presented on a visual display. Eliminating any visual input we demonstrate that alpha power modulations associated with auditory cueing indeed encipher physical space. Previous reports suggest that anticipation of auditory distractors already involves lateralization of posterior alpha power [9] and that this lateralization scales monotonically with the interaural time difference. That is, the subjectively perceived lateral position[12]. In line with these observations, present analyses strengthen the notion that power modulations of alpha-band activity reflect a mechanism by which space and direction are inferred by the observer. Thus, we propose that alpha power modulation is egocentric and it reflects the participant's directional heading.

Alpha power modulation allows decoding of auditory covert attention

In visual spatial attention, a large body of evidence suggests that the direction of attention can be decoded on the basis of posterior alpha activity [14,20,40-42] using forward encoding models[43]. Here we confirm that this finding generalizes to the auditory domain and extend to directions beyond the visual field (i.e. to the sides and behind the participant). That is, posterior alpha power modulation does not simply reflect suppression of anticipated interfering visual input. Instead, it reflects an active process of tuning to sound origin and directing attention to optimize performance (e.g. faster reaction times correlated with stronger tuning Figure 4). To what extend this tuning is specific to alpha oscillations merits further examination. In the present report, the signal-to-noise ratio in the other frequency bands (e.g. theta or gamma) is relatively poor. Foster et al. reported that the neural tuning to a spatial location, in fact tracks

the spatial representation held in WM [40]. This tuning was evident for the alpha band but absent for theta and gamma-band activity. The authors noted that the distribution of activity in the latter frequency bands could lack location specificity yet still be relevant for the coding of the spatial aspects in the context of the WM task. In conclusion, based on the present analyses, tuning effects to location manifesting in other frequency bands cannot be excluded. Nonetheless, decoding is strongest in the alpha band.

Alpha power lateralization corresponds to lateralization of slow ERPs

Motivated by the nearly identical spatial distribution of the lateralization of the auditory ERPs (Figure 2B) and ongoing alpha activity (Figure 3B) we conducted an exploratory analysis to establish a relationship between the two. Indeed, their time courses were remarkably similar (Figure 3C). The strongest correlation was observed around 400ms post cue onset corresponding to both the strongest alpha power modulation (decrease relative to pre-cue baseline) and strongest negativity of slow auditory ERPs (Figure 2A). Theoretical and empirical work suggests that this relationship can be understood as a consequence of the asymmetric amplitude modulation of ongoing alpha activity[30-32]. This non-sinusoidal shape of neuronal oscillations, in general is not a mere phenomenological nuisance but an important physiological feature to be considered in the quest to understand neural communication and the emergence of cognition[44,45]. Briefly, as a consequence of the asymmetry peaks is modulated stronger than throughs (or *vice versa* depending on the reference), stimulus events induce a power modulation that does not fluctuate around zero. Instead, is characterized by slow sustained potential following trial averaging and manifesting as slow event-related potential. Thus, slow ERP's and cue-induced modulation of ongoing alpha activity potentially reflect the two sides of the same coin. As the former is obtained through averaging across repetitions it is less suited for decoding of spatial location. Nonetheless, present results

highlight the relevance of the parieto-occipital cortex in tracking spatial sound origin in audition beyond the mere allocation of attention and interference suppression.

Impulses for the multimodal view of the brain and the dual processing stream theory

Present results open novel empirical questions both in the fields of visual and auditory attention but also directed towards our current understanding of the multimodal brain.

The involvement of saccades in register with the attended loudspeaker location provides evidence for the existence of a reciprocal relationship as the recently discovered saccades induced eardrum oscillations [36,37]. That is, an auditory cue presentation at a particular location in space elicits oculomotor responses consistent with the sound origin. A key structure in generating saccades in the superior colliculus (SC)[46]. It has been demonstrated that subcortical auditory nuclei in the brainstem project to the SC enabling the instantiation of an auditory space map [47]. In light of the present results that alpha power lateralization known from visual attention studies is instantiated during purely a auditory task, an emerging question is to what extend alpha power modulation in audition reflects an active orienting response following saccades? What is the role of this saccadic behavior in shaping attention direction both in vision and audition? A study by Wöstmann et al. studied auditory distractor suppression and found that alpha power lateralization over occipital regions informs the suppression of the interfering auditory input[9]. Alpha power was increased contralateral to distractor location. The authors reasoned that indeed, saccadic behavior might influence the conclusions and performed a saccadic localizer task prior to the auditory experiment. Noteworthy, artifact control of eye movements eliminates the muscular contribution to the EEG scalp topography, yet the consequence of the movement registered by the sensory system remains unaltered, potentially manifesting in the modulation of posterior alpha activity.

The present findings provide empirical support for the auditory dorsal stream model as an internal model for sensorimotor control[48]. According to this model, the operational characteristic of the dorsal stream is understood as an "internal model" of the surrounding environment relative to the individual's vantage point. It implements sensorimotor sequences from all sensory modalities into a unified experience in space-time by combining elements of 'where', 'when', and 'how' in order to guide coordinated behavior. This notion requires at least a 1) orientation to a sound source in the external world that is consistent with the motor output (e.g. saccade direction) of the observing individual and 2) gain increase in the sensory system to support processing (e.g. alpha power modulation). We have provided evidence for both requirements. We argue that the modulation of alpha activity reflects an observer related increase in gain initiated by the spatial orientation towards the sound origin together with a consistent motor response [37,49].

In conclusion, present results confirm a multimodal functional relevance of alpha oscillatory activity that reflects the integration of auditory and visual utilities of the observing individual into a direction-specific sensory gain increase to organize and instantiate coordinated behavior.

Acknowledgments: This work was supported by the James S. McDonnel Foundation Understanding Human Cognition Collaborative Award 220020448. The authors thank all participants volunteering in this study.

Conflict of interest: The authors declare no competing financial interests.

References

1 Kelly, S. P., Lalor, E. C., Reilly, R. B. & Foxe, J. J. Increases in alpha oscillatory power reflect an active retinotopic mechanism for distracter suppression during sustained visuospatial attention. *J Neurophysiol* **95**, 3844-3851, doi:10.1152/jn.01234.2005 (2006).

2 Popov, T., Gips, B., Kastner, S. & Jensen, O. Spatial specificity of alpha oscillations in the human visual system. *Hum Brain Mapp* **40**, 4432-4440, doi:10.1002/hbm.24712 (2019).

3 Rihs, T. A., Michel, C. M. & Thut, G. Mechanisms of selective inhibition in visual spatial attention are indexed by alpha-band EEG synchronization. *The European journal of neuroscience* **25**, 603-610, doi:10.1111/j.1460-9568.2007.05278.x (2007).

4 Michalka, S. W., Rosen, M. L., Kong, L., Shinn-Cunningham, B. G. & Somers, D. C. Auditory Spatial Coding Flexibly Recruits Anterior, but Not Posterior, Visuotopic Parietal Cortex. *Cereb Cortex* **26**, 1302-1308, doi:10.1093/cercor/bhv303 (2016).

5 Rauschecker, J. P. & Tian, B. Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proc Natl Acad Sci U S A* **97**, 11800-11806, doi:10.1073/pnas.97.22.11800 (2000).

6 van der Heijden, K., Rauschecker, J. P., de Gelder, B. & Formisano, E. Cortical mechanisms of spatial hearing. *Nat Rev Neurosci* **20**, 609-623, doi:10.1038/s41583-019-0206-5 (2019).

7 Klatt, L. I., Getzmann, S., Wascher, E. & Schneider, D. Searching for auditory targets in external space and in working memory: Electrophysiological mechanisms underlying perceptual and retroactive spatial attention. *Behav Brain Res* **353**, 98-107, doi:10.1016/j.bbr.2018.06.022 (2018).

8 Muller, N. & Weisz, N. Lateralized auditory cortical alpha band activity and interregional connectivity pattern reflect anticipation of target sounds. *Cereb Cortex* **22**, 1604-1613, doi:10.1093/cercor/bhr232 (2012).

9 Wostmann, M., Alavash, M. & Obleser, J. Alpha Oscillations in the Human Brain Implement Distractor Suppression Independent of Target Selection. *J Neurosci* **39**, 9797-9805, doi:10.1523/JNEUROSCI.1954-19.2019 (2019).

10 Wostmann, M., Herrmann, B., Maess, B. & Obleser, J. Spatiotemporal dynamics of auditory attention synchronize with speech. *Proc Natl Acad Sci U S A* **113**, 3873-3878, doi:10.1073/pnas.1523357113 (2016).

11 Kerlin, J. R., Shahin, A. J. & Miller, L. M. Attentional gain control of ongoing cortical speech representations in a "cocktail party". *J Neurosci* **30**, 620-628, doi:10.1523/JNEUROSCI.3631-09.2010 (2010).

12 Deng, Y., Choi, I. & Shinn-Cunningham, B. Topographic specificity of alpha power during auditory spatial attention. *Neuroimage* **207**, 116360, doi:10.1016/j.neuroimage.2019.116360 (2020).

13 Tune, S., Wostmann, M. & Obleser, J. Probing the limits of alpha power lateralisation as a neural marker of selective attention in middle-aged and older listeners. *The European journal of neuroscience* **48**, 2537-2550, doi:10.1111/ejn.13862 (2018).

14 Foster, J. J., Bsales, E. M., Jaffe, R. J. & Awh, E. Alpha-Band Activity Reveals Spontaneous Representations of Spatial Position in Visual Working Memory. *Curr Biol* **27**, 3216-3223 e3216, doi:10.1016/j.cub.2017.09.031 (2017).

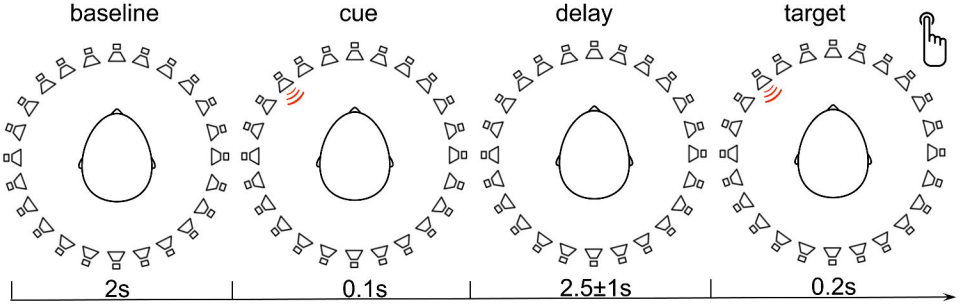
15 Waldhauser, G. T., Braun, V. & Hanslmayr, S. Episodic Memory Retrieval Functionally Relies on Very Rapid Reactivation of Sensory Information. *J Neurosci* **36**, 251-260, doi:10.1523/JNEUROSCI.2101-15.2016 (2016).
16 Hamm, J. P., Sabatinelli, D. & Clementz, B. A. Alpha oscillations and the control of voluntary saccadic behavior. *Exp Brain Res* **221**, 123-128, doi:10.1007/s00221-012-3167-8 (2012).
17 Ito, J., Maldonado, P., Singer, W. & Grun, S. Saccade-related modulations of neuronal excitability support synchrony of visually elicited spikes. *Cereb Cortex* **21**, 2482-2497, doi:10.1093/cercor/bhr020 (2011).
18 Lowet, E. *et al*. Microsaccade-rhythmic modulation of neural synchronization and coding within and across cortical areas V1 and V2. *PLoS Biol* **16**, e2004132, doi:10.1371/journal.pbio.2004132 (2018).
19 Lowet, E. *et al*. Enhanced Neural Processing by Covert Attention only during Microsaccades Directed toward the Attended Stimulus. *Neuron* **99**, 207-214 e203, doi:10.1016/j.neuron.2018.05.041 (2018).
20 Popov, T., Kastner, S. & Jensen, O. FEF-Controlled Alpha Delay Activity Precedes Stimulus-Induced Gamma-Band Activity in Visual Cortex. *J Neurosci* **37**, 4117-4127, doi:10.1523/JNEUROSCI.3015-16.2017 (2017).
21 Oostenveld, R., Fries, P., Maris, E. & Schoffelen, J. M. FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci* **2011**, 156869, doi:10.1155/2011/156869 (2011).
22 Jung, T. P. *et al*. Imaging Brain Dynamics Using Independent Component Analysis. *Proc IEEE Inst Electr Electron Eng* **89**, 1107-1122, doi:10.1109/5.939827 (2001).
23 Gross, J. *et al*. Dynamic imaging of coherent sources: Studying neural interactions in the human brain. *Proc Natl Acad Sci U S A* **98**, 694-699, doi:10.1073/pnas.98.2.694 (2001).
24 Gramfort, A., Papadopoulo, T., Olivi, E. & Clerc, M. OpenMEEG: opensource software for quasistatic bioelectromagnetics. *Biomed Eng Online* **9**, 45, doi:10.1186/1475-925X-9-45 (2010).
25 Desikan, R. S. *et al*. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* **31**, 968-980, doi:10.1016/j.neuroimage.2006.01.021 (2006).
26 Van Veen, B. D., van Drongelen, W., Yuchtman, M. & Suzuki, A. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Trans Biomed Eng* **44**, 867-880, doi:10.1109/10.623056 (1997).
27 Haufe, S. *et al*. On the interpretation of weight vectors of linear models in multivariate neuroimaging. *Neuroimage* **87**, 96-110, doi:10.1016/j.neuroimage.2013.10.067 (2014).
28 Maris, E. & Oostenveld, R. Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* **164**, 177-190, doi:10.1016/j.jneumeth.2007.03.024 (2007).
29 Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R. & Kievit, R. A. Raincloud plots: a multi-platform tool for robust data visualization. *Wellcome Open Res* **4**, 63-63, doi:10.12688/wellcomeopenres.15191.1 (2019).
30 Jensen, O., van Dijk, H. & Mazaheri, A. Amplitude asymmetry as a mechanism for the generation of slow evoked responses. *Clin Neurophysiol* **121**, 1148-1149; author reply 1149-1150, doi:10.1016/j.clinph.2010.01.037 (2010).
31 van Dijk, H., van der Werf, J., Mazaheri, A., Medendorp, W. P. & Jensen, O. Modulations in oscillatory activity with amplitude asymmetry can produce cognitively relevant event-related responses. *Proc Natl Acad Sci U S A* **107**, 900-905, doi:10.1073/pnas.0908821107 (2010).

32 Iemi, L. *et al*. Multiple mechanisms link prestimulus neural oscillations to sensory responses. *Elife* **8**, doi:10.7554/eLife.43620 (2019).
33 Adam, K. C. S., Robison, M. K. & Vogel, E. K. Contralateral Delay Activity Tracks Fluctuations in Working Memory Performance. *J Cogn Neurosci* **30**, 1229-1240, doi:10.1162/jocn_a_01233 (2018).
34 Cohen, Y. E., Russ, B. E. & Gifford, G. W., 3rd. Auditory processing in the posterior parietal cortex. *Behav Cogn Neurosci Rev* **4**, 218-231, doi:10.1177/1534582305285861 (2005).
35 Braga, R. M., Fu, R. Z., Seemungal, B. M., Wise, R. J. & Leech, R. Eye Movements during Auditory Attention Predict Individual Differences in Dorsal Attention Network Activity. *Front Hum Neurosci* **10**, 164, doi:10.3389/fnhum.2016.00164 (2016).
36 Murphy, D. L., King, C. D., Schlebusch, S. N., Shera, C. A. & Groh, J. M. Evidence for a system in the auditory periphery that may contribute to linking sounds and images in space. *bioRxiv*, 2020.2007.2019.210864, doi:10.1101/2020.07.19.210864 (2020).
37 Gruters, K. G. *et al*. The eardrums move when the eyes move: A multisensory effect on the mechanics of hearing. *Proceedings of the National Academy of Sciences* **115**, E1309-E1318, doi:10.1073/pnas.1717948115 (2018).
38 Frey, J. N., Ruhnau, P. & Weisz, N. Not so different after all: The same oscillatory processes support different types of attention. *Brain research* **1626**, 183-197, doi:10.1016/j.brainres.2015.02.017 (2015).
39 Fu, K. M. *et al*. Attention-dependent suppression of distracter visual input can be cross-modally cued as indexed by anticipatory parieto-occipital alpha-band oscillations. *Brain Res Cogn Brain Res* **12**, 145-152, doi:10.1016/s0926-6410(01)00034-9 (2001).
40 Foster, J. J., Sutterer, D. W., Serences, J. T., Vogel, E. K. & Awh, E. The topography of alpha-band activity tracks the content of spatial working memory. *J Neurophysiol* **115**, 168-177, doi:10.1152/jn.00860.2015 (2016).
41 Samaha, J., Sprague, T. C. & Postle, B. R. Decoding and Reconstructing the Focus of Spatial Attention from the Topography of Alpha-band Oscillations. *J Cogn Neurosci* **28**, 1090-1097, doi:10.1162/jocn_a_00955 (2016).
42 Munneke, J., Fahrenfort, J., Sutterer, D., Theeuwes, J. & Awh, E. Multivariate analysis of EEG activity indexes contingent and non-contingent attentional capture. *bioRxiv*, 734004, doi:10.1101/734004 (2019).
43 Brouwer, G. J. & Heeger, D. J. Decoding and reconstructing color from responses in human visual cortex. *J Neurosci* **29**, 13992-14003, doi:10.1523/JNEUROSCI.3577-09.2009 (2009).
44 Cole, S. R. & Voytek, B. Brain Oscillations and the Importance of Waveform Shape. *Trends Cogn Sci* **21**, 137-149, doi:10.1016/j.tics.2016.12.008 (2017).
45 Jackson, N., Cole, S. R., Voytek, B. & Swann, N. C. Characteristics of Waveform Shape in Parkinson's Disease Detected with Scalp Electroencephalography. *eNeuro* **6**, doi:10.1523/ENEURO.0151-19.2019 (2019).
46 Schiller, P. H. in *Comprehensive Physiology* 457-505 (2011).
47 Mellott, J. G., Beebe, N. L. & Schofield, B. R. GABAergic and non-GABAergic projections to the superior colliculus from the auditory brainstem. *Brain Struct Funct* **223**, 1923-1936, doi:10.1007/s00429-017-1599-4 (2018).
48 Rauschecker, J. P. Where, When, and How: Are they all sensorimotor? Towards a unified view of the dorsal pathway in vision and audition. *Cortex* **98**, 262-268, doi:10.1016/j.cortex.2017.10.020 (2018).

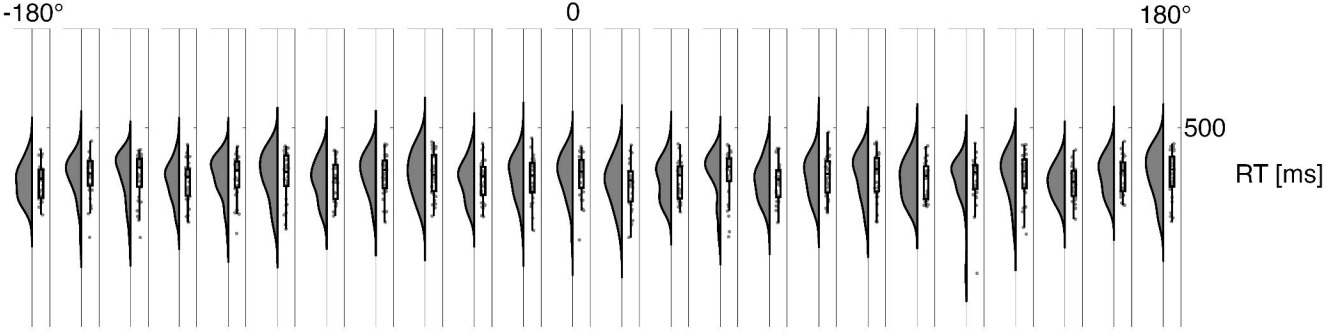
49 Porter, K. K., Metzger, R. R. & Groh, J. M. Visual- and saccade-related signals in the primate inferior colliculus. *Proceedings of the National Academy of Sciences* **104**, 17855-17860, doi:10.1073/pnas.0706249104 (2007).

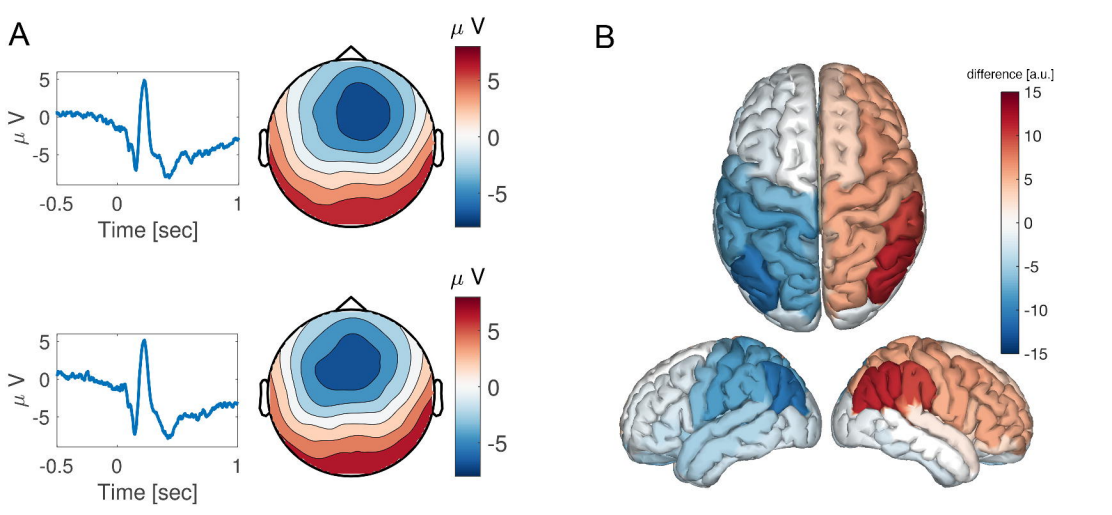
A

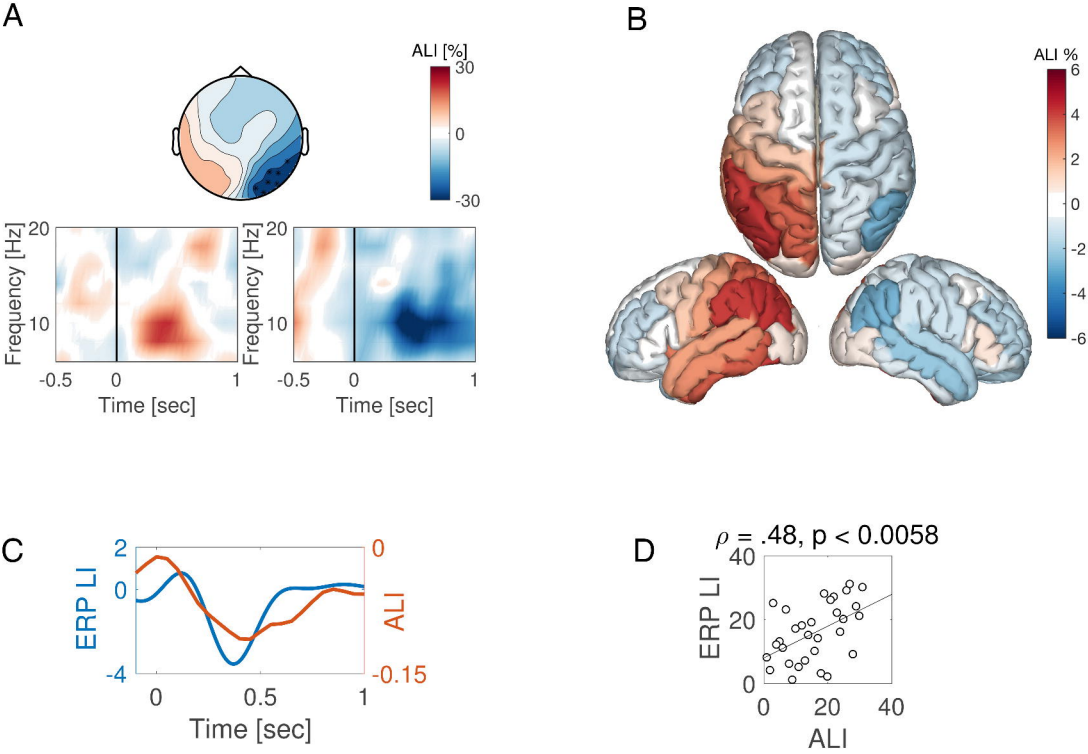
B

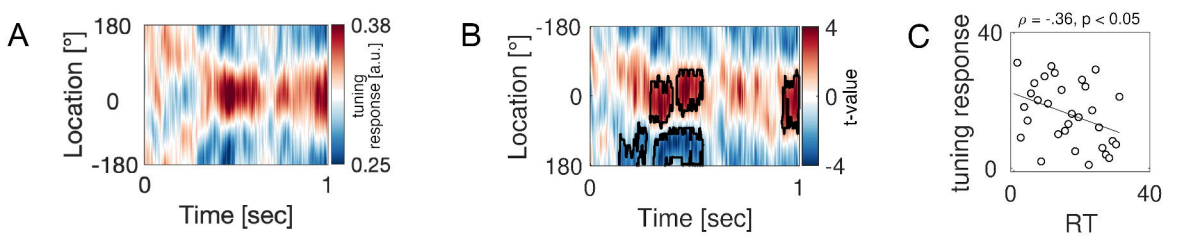

baseline
cue
delay
target
2s
0.1s
2.5±1s
0.2s

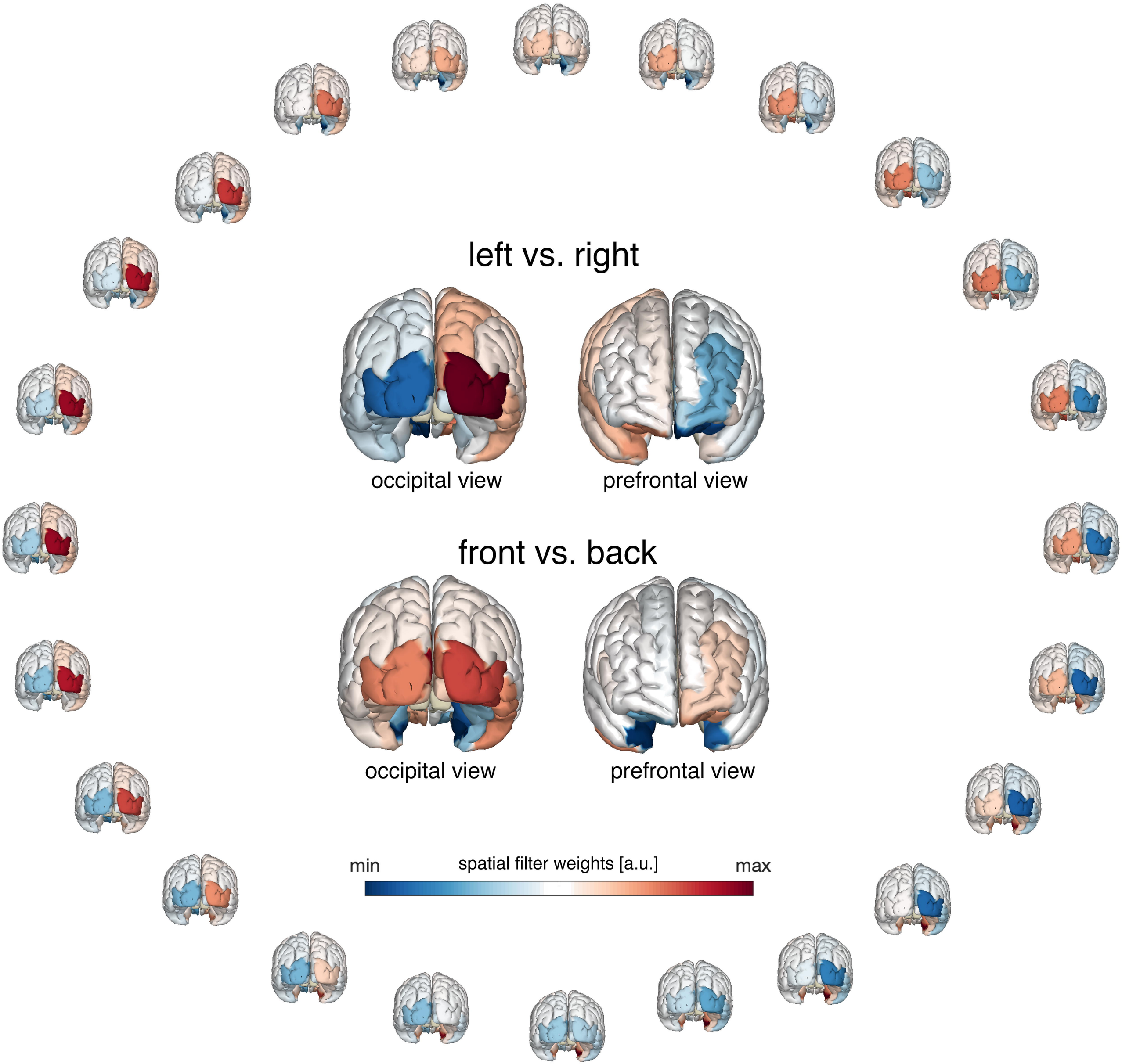
C

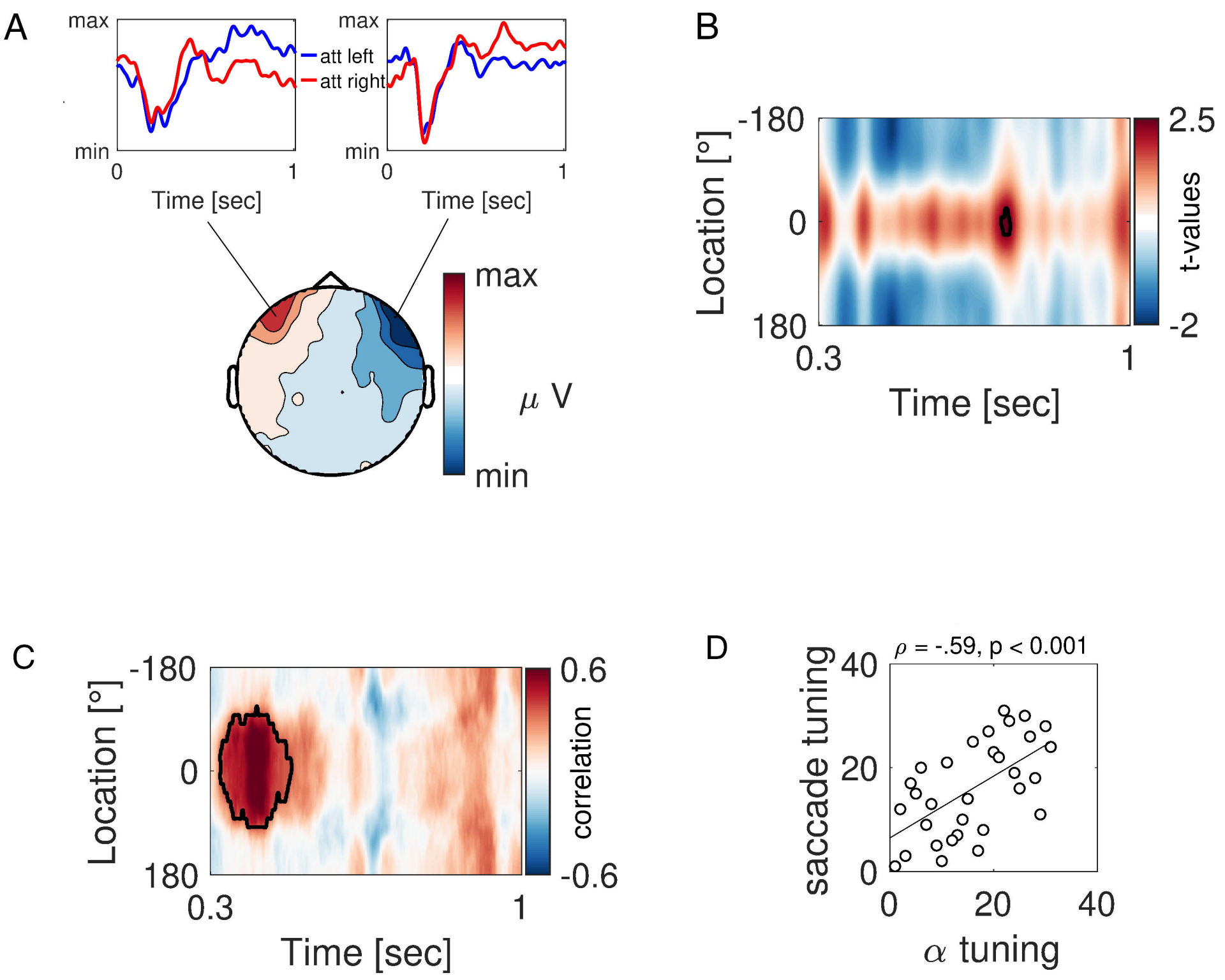

Loudspeaker angle [°]
-180°
0
180°
500
RT [ms]


A
μ V
5
0
-5
-0.5
0
1
Time [sec]
μ V
5
0
-5
μ V
5
0
-5
-0.5
0
1
Time [sec]
μ V
5
0
-5
B
difference [a.u.]
15
10
5
0
-5
-10
-15


A
ALI [%]
30
0
-30
Frequency [Hz]
20
10
-0.5
0
1
Time [sec]
B
ALI %
6
4
2
0
-2
-4
-6
C
ERP LI
ALI
2
0
-4
0
-0.15
0
0.5
1
Time [sec]
D
ρ = .48, p < 0.0058
ERP LI
ALI
40
20
0


A
180
0
-180
Location [°]
0
1
Time [sec]
0.38
tuning
response [a.u.]
0.25
B
-180
0
180
Location [°]
0
1
Time [sec]
4
0
-4
t-value
C
ρ = -.36, p < 0.05
40
0
tuning response
0
40
RT


left vs. right
occipital view
prefrontal view
front vs. back
occipital view
prefrontal view
min
spatial filter weights [a.u.]
max


A
max
min
0
1
Time [sec]
att left
att right
max
min
0
1
Time [sec]
max
μ V
min
B
Location [°]
-180
0
180
0.3
1
Time [sec]
2.5
t-values
-2
C
Location [°]
-180
0
180
0.3
1
Time [sec]
0.6
correlation
-0.6
D
ρ = -.59, p < 0.001
saccade tuning
40
20
0
0
20
40
α tuning