Spatiotemporal dynamics of auditory attention synchronize with speech

Malte Wöstmann[a,b,1], Björn Herrmann[b,c], Burkhard Maess[d], and Jonas Obleser[a,b,1]

[a]Department of Psychology, University of Lübeck, 23562 Luebeck, Germany; [b]Max Planck Research Group "Auditory Cognition," Max Planck Institute for Human Cognitive and Brain Sciences, 04103 Leipzig, Germany; [c]Department of Psychology, The University of Western Ontario, London, Canada N6A 5C2; and [d]Magnetoencephalography and Cortical Networks Unit, Max Planck Institute for Human Cognitive and Brain Sciences, 04103 Leipzig, Germany

Attention plays a fundamental role in selectively processing stimuli in our environment despite distraction. Spatial attention induces increasing and decreasing power of neural alpha oscillations (8–12 Hz) in brain regions ipsilateral and contralateral to the locus of attention, respectively. This study tested whether the hemispheric lateralization of alpha power codes not just the spatial location but also the temporal structure of the stimulus. Participants attended to spoken digits presented to one ear and ignored tightly synchronized distracting digits presented to the other ear. In the magnetoencephalogram, spatial attention induced lateralization of alpha power in parietal, but notably also in auditory cortical regions. This alpha power lateralization was not maintained steadily but fluctuated in synchrony with the speech rate and lagged the time course of low-frequency (1–5 Hz) sensory synchronization. Higher amplitude of alpha power modulation at the speech rate was predictive of a listener's enhanced performance of stream-specific speech comprehension. Our findings demonstrate that alpha power lateralization is modulated in tune with the sensory input and acts as a spatiotemporal filter controlling the read-out of sensory content.

attention | neural oscillations | alpha lateralization | synchronization | speech

Neural oscillations are tenable biological substrates to instantiate attentional selection of sensory information from our environment (1–3). The power of alpha oscillations (8–12 Hz) is modulated by the degree of selective attention to a behaviorally relevant stimulus (4–6). Attentive focusing to one side in auditory, visual, or tactile space leads to a relative decrease in alpha power in contralateral compared with ipsilateral sensory brain regions (7–10) and governs success of selective attention, that is, of isolating one stimulus at a specific spatial location (11–13) in the context of other distracting stimuli. In speech, however, spatial selective attention alone does not suffice for successful speech comprehension given the multiple timescales (e.g., syllable and word rate) over which speech varies. Any attentional neural mechanism that respects the temporal structure of auditory sensory inputs must therefore be dynamic over time. Here, we tested the hypothesis that spatial attention to one of two concurrent speech streams induces a lateralization of alpha power that synchronizes with the speech rate over time and enhances speech comprehension.

In addition to neural alpha oscillations and their role in selective attention, low-frequency neural oscillations (delta/theta band; 1–5 Hz) in sensory cortices synchronize with the temporal structure of acoustic stimuli (14, 15) such as human speech (16, 17), and this synchronization supports perception (18). Furthermore, synchronization of low-frequency neural oscillations with speech is enhanced for speech streams to which participants attend (19, 20). Critically, low-frequency oscillations are commonly specified as phase-locked activity (synchronizing with acoustic events) whereas alpha oscillations are quantified as non-phase–locked activity, and both have thus been often investigated in isolation. If we are, however, to understand the human ability to flexibly track speech over time in a multitalker situation we must understand how low-frequency,

sensory-driven neural activity and top-down, attention-related alpha oscillations interplay to influence speech comprehension (21).

Using an attentive-listening design with concurrent, tightly synchronized dichotic input, the current magnetoencephalography (MEG) study shows that alpha power lateralizes with respect to the side of attention and that alpha power lateralization synchronizes with the temporal structure of speech time-delayed relative to low-frequency sensory synchronization. Critically, the magnitude of these temporal modulations of alpha power predicts successful recall of to-be-attended stimuli.

Results

Participants ($n = 19$) were asked to attend to four spoken digits presented to one ear (indicated by a monaurally presented cue tone) and to ignore four different digits presented concurrently and perceptually onset-aligned (22) to the other ear (Fig. 1A). Participants were asked to report the four digits that had been presented on the to-be-attended side by selecting them from a visually presented 12-digit array. The correct selection of a digit presented to the to-be-attended side was considered a "hit." Incorrect selections of digits were split into "spatial confusions" (the selected digit had occurred on the to-be-ignored side) and "random errors" (the selected digit had not been presented).

Ignored Speech Causes Confusion with Attended Speech. Fig. 1B shows proportions of hits, spatial confusions, and random errors that differed significantly (Friedman test; $\chi^2 (2) = 35.71$; $P < 0.001$; $r = 0.94$). The proportion of hits was higher than spatial confusions (Wilcoxon signed-rank test; $z = 3.82$; $P < 0.001$; $r = $

Significance

Attending and ignoring play fundamental roles in our everyday behavior in spatially and temporally fast-varying environments. Does focused attention to a specific stimulus adapt to these spatiotemporal dynamics? Using magnetoencephalography, we investigated the differential time courses of sensory encoding of two simultaneous, bottom-up rhythmic speech streams versus neural signatures of top-down spatial attention to one of the two streams. As expected, spatial attention modulated the power of ~10-Hz (alpha) oscillations and their lateralization across the two cerebral hemispheres. Critically, this lateralization of alpha activity was modulated in temporal synchrony with the speech rate and predicted participants' speech comprehension. Our results demonstrate that alpha activity acts as a spatiotemporal filter to control the read-out of attended and ignored sensory content.
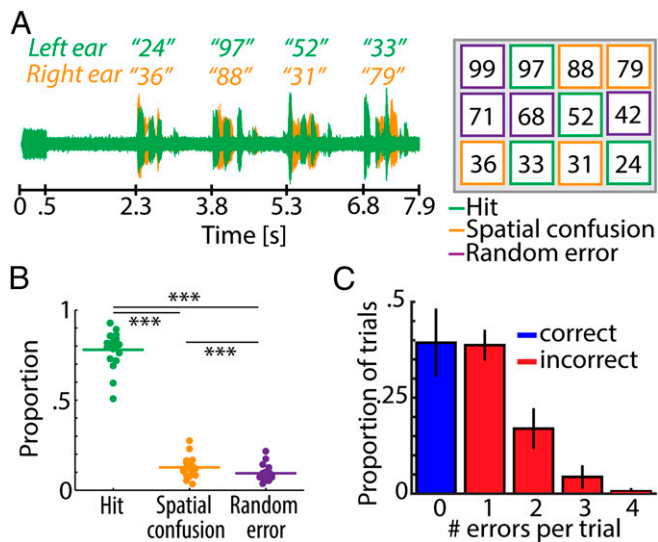

NEUROSCIENCE

PNAS

Fig. 1. (A) Trial design. An auditory spatial cue (1,000-Hz pure tone; 500 ms; left or right ear) indicated the to-be-attended side. After stimulus anticipation (0.5–2.3 s), four spoken digits were presented to the left and four different digits to the right ear (2.3–7.9 s). Auditory materials were presented in random noise (+10 dB signal-to-noise ratio). Participants were asked to select the four digits that were presented on the to-be-attended side from a visually presented array of digits. Each selected digit could either be a hit (selected digit appeared on to-be-attended side; green), a spatial confusion (selected digit appeared on to-be-ignored side; orange) or a random error (selected digit was not presented; purple; colored edges were not shown in the actual experiment). (B) Dots show the individual participants' proportions of response types (n = 19). Horizontal lines show the mean across participants. ***P < 0.001. (C) Bars show mean proportions of trials as a function of the number of errors on a trial (0–4). Trials without any errors were classified as correct, trials with one or more errors were classified as incorrect. Error bars show 95% confidence intervals.

0.97) and random errors ($z = 3.82$; $P < 0.001$; $r = 0.98$). Critically, the proportion of spatial confusions was higher than the proportion of random errors ($z = 3.68$; $P < 0.001$; $r = 0.31$), indicating that to-be-ignored digits interfered with the recall of to-be-attended digits (for further analyses of behavioral results, see Fig. S1).

Low-Frequency Phase Coherence Versus Alpha Power. We investigated two neural signatures in the MEG: phase coherence of low-frequency oscillations [quantified as 1- to 5-Hz intertrial phase coherence (ITPC); Fig. 2 A and B] and the power of alpha oscillations (8–12 Hz; Fig. 2 C and D). Phase coherence was tightly time-locked to the acoustic stimulation, with peaks of phase coherence occurring at the cue on- and offset and at digit onsets. Neural generators were localized in the bilateral superior temporal cortex. In contrast, alpha

oscillatory power was strongest before trial onset and decreased gradually throughout the trial, with the main generators localized in the occipital lobe.

Note that high overall (i.e., condition-independent) alpha power (Fig. 2 C and D) indicates suppression of neural activity in task-irrelevant visual areas in the occipital lobe (23–25), which in turn might support enhanced neural function in the task-relevant auditory modality (26). In the current study, high alpha power shortly before trial onset (0 s) might support attentional focusing to the auditory cue, which was indispensable for realizing to which side (left vs. right) the participant was supposed to attend.

Task-Induced Hemispheric Lateralization of Phase Coherence and Alpha Power. We tested the effect of attention conditions (left vs. right) on the topographies of neural responses and calculated the Attentional Modulation Index [$AMI_X = (X_{att\_left} - X_{att\_right})/(X_{att\_left} + X_{att\_right})$] with X representing either phase coherence (AMI_{ITPC}) or alpha power (AMI_α) at all MEG sensors (8). A positive AMI indicates larger neural responses for attention-left trials and negative AMI larger responses for attention-right trials. A difference of the AMI between the left and the right hemisphere is evidence for a hemispheric lateralization of neural responses by the spatial attention demands under a given condition.

Mean AMI_{ITPC} was higher at sensors on the right compared with the left hemisphere during cue presentation (Fig. 3A; 0–0.5 s; Wilcoxon signed-rank test; $z = 3.34$; $P < 0.001$; $r = 0.56$) and stimulus anticipation (0.5–2.3 s; $t_{18} = 2.16$; $P = 0.045$; $r = 0.24$), but not during speech stimulus presentation (2.3–7.9 s; $t_{18} = 0.32$; $P = 0.75$; $r = 0.03$). To the contrary, AMI_α was positive at sensors over the left and negative at sensors over the right hemisphere (Fig. 3B). The hemispheric difference of AMI_α not only was significant during cue presentation (Wilcoxon signed-rank test; $z = 2.13$; $P = 0.033$; $r = 0.21$) and stimulus anticipation ($z = 2.82$; $P = 0.005$; $r = 0.27$), but also was sustained during speech stimulus presentation ($t_{18} = 2.55$; $P = 0.02$; $r = 0.14$). Note that the lateralization of neural oscillatory power was significant only for frequencies in the alpha band (Fig. S2).

Attention Modulates Alpha Power in Parietal and Auditory Cortex Regions. For each participant, beamformer source reconstruction (27) was used to calculate the spatial distribution of AMI_α (alpha power lateralization). Fig. 4 shows brain surfaces overlaid with AMI_α values that were significantly different from zero during speech stimulus presentation (2.3–7.9 s; t test; $P < 0.05$; uncorrected). AMI_α was positive in left occipital regions and negative in right inferior parietal and inferior frontal regions. Critically, significant AMI_α values were found in auditory cortex regions in the left ($AMI_\alpha > 0$) and even stronger in the right hemisphere ($AMI_\alpha < 0$; see also Fig. S3 for source reconstruction of AMI_α in cue and anticipation periods).
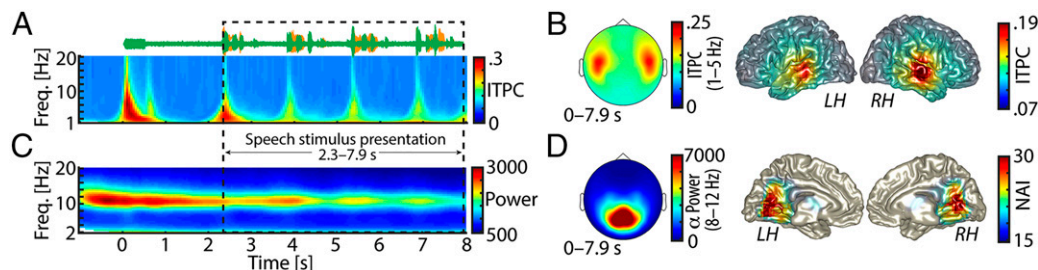
Fig. 2. (A) ITPC and (C) power of induced oscillations in the dichotic listening task averaged across all trials (attention-left and attention-right), 102 combined gradiometer sensors, and 19 participants. Note strong ITPC in low frequencies (1–5 Hz) time-locked to the acoustic stimulation and high power of alpha oscillations (8–12 Hz) throughout the trial. (B) Beamformer source reconstructions revealed posterior temporal cortex regions as the major generators of phase coherence and (D) occipital regions as generators for alpha power.

A Lateralization of 1–5 Hz phase coherence
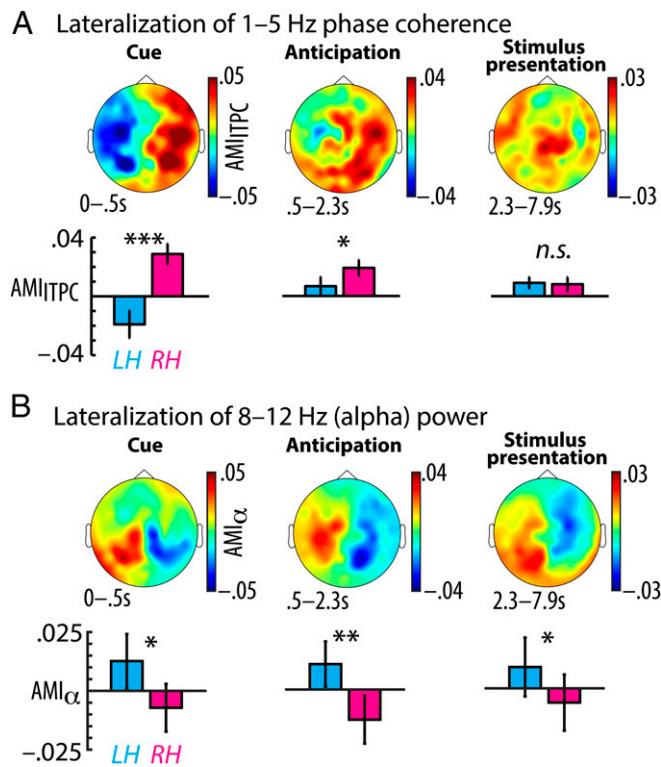
Fig. 3. Topographic maps of the attentional modulation index in three time periods (cue, anticipation, speech stimulus presentation) for low-frequency ITPC (1–5 Hz AMI_ITPC; *A, Top*) and alpha power (8–12 Hz AMI_α; *B, Top*). Bar graphs show mean across all sensors on the left hemisphere (LH) and right hemisphere (RH) for AMI_ITPC (*A, Bottom*) and AMI_α (*B, Bottom*). Error bars indicate ±1 SEM. AMI_ITPC showed a significant hemispheric lateralization (RH > LH) only early during a trial, before presentation of the speech stimulus. The hemispheric difference in AMI_α showed the opposite pattern (LH > RH) and was significant during the entire trial including speech stimulus presentation. *P < 0.05; **P < 0.01; ***P < 0.001; n.s., not significant.

Lateralized Alpha Power Is Modulated in Synchrony with the Speech Rate and Predicts Recall of To-Be-Attended Speech. For each participant, we selected 20 MEG sensors with the largest positive AMI_α values on the left hemisphere and 20 sensors with the largest negative AMI_α values on the right hemisphere (Figs. S4 and S5A). To obtain a time-resolved metric of the attentional modulation of oscillatory power, alpha power at ipsilateral sensors α_{att_ipsi} (i.e., ipsilateral to the to-be-attended side) and contralateral sensors α_{att_contra} (i.e., contralateral to the to-be-attended side) were contrasted to calculate the Alpha Lateralization Index [ALI = $(\alpha_{att\_ipsi} - \alpha_{att\_contra})/(\alpha_{att\_ipsi} + \alpha_{att\_contra})$] (8). Fig. 5 shows time courses of the ALI as well as the low-frequency phase coherence (1–5 Hz) superimposed on a stimulus waveform (for time courses of α_{att_ipsi} and α_{att_contra}, see Fig. S6). During speech stimulus presentation (2.3–7.9 s), digits were presented at a word rate of 0.67 Hz, and a phase coherence peak instantly followed each digit onset. Critically, ALI also showed characteristic modulations, with ALI peaks co-occurring with midpoints of spoken digits and thus lagging behind peaks of phase coherence (temporal modulation of the ALI was driven to equal extends by ipsilateral and contralateral alpha power; Fig. S7). The 0.67-Hz phase delay between phase coherence and ALI was significant (Parametric Hotelling paired-sample test; mean delay: 2.4 rad; $F_{34} = 116.19$; $P < 0.001$; $r = 0.93$; Fig. 5B), and ALI and ITPC were significantly phase-concentrated across participants (circular Rayleigh test; ALI: $z = 5.55$; $P = 0.003$; $r = 0.54$; ITPC: $z = 17.78$; $P < 0.001$; $r = 0.97$).

We tested whether the amplitude modulations of the ALI predict the recall of to-be-attended digits. To this end, trials with only correct responses were classified as "correct" and trials with one or more incorrect responses as "incorrect" (Fig. 1C). Fig. 6A shows amplitude spectra (from Fast Fourier Transforms) of the ALI during speech stimulus presentation (2.3–7.9 s) for correct and incorrect trials (for ALI time courses of correct and incorrect trials, see Fig. S8). The amplitude of 0.67-Hz modulations of ALI was significantly higher in correct compared with incorrect trials ($t_{18} = 2.63$; $P = 0.017$; $r = 0.38$; no significant 0.67-Hz phase difference between correct and incorrect trials: $F_{34} = 1.12$; $P = 0.348$; $r = 0.26$). There was no difference of 1- to 5-Hz phase coherence modulations at 0.67 Hz between correct and incorrect trials (Fig. S9).

Furthermore, in incorrect trials, participants with higher amplitudes in their 0.67-Hz modulations of ALI made fewer errors (Fig. 6B; $r_{Spearman} = -0.58$; $P = 0.01$). Note that there was no significant circular–linear correlation between the number of errors and the 0.67-Hz ALI phase ($r = 0.11$; $P = 0.48$) or phase distance between ITPC and ALI ($r = 0.08$; $P = 0.7$), respectively.

Discussion

This study shows that neural alpha oscillations act as a spatially selective and temporally synchronized filter of attention during speech comprehension. Specifically, attention to one of two dichotically presented streams of spoken digits increased alpha power in supramodal parietal and sensory temporal cortex regions in the hemisphere ipsilateral to the attended stream (and decreased alpha power contralaterally). Importantly, the hemispheric lateralization of alpha power was (*i*) modulated in synchrony with the speech rate, (*ii*) lagged behind the time course of the low-frequency phase coherence related to sensory encoding (1–5 Hz), and (*iii*) predicted accurate recall of to-be-attended stimuli.

Spatial Alpha Power Modulation Reveals Dynamics of Selective Attention over Time. In our dichotic listening task, increases in low-frequency, phase-locked neural activity (1–5 Hz) coincided with auditory events over time (Fig. 2A). Localization of phase-locked activity demonstrates that the auditory cortex reliably tracks the two concurrent streams of speech (16, 17, 28). However, phase-locked neural activity did not show any preference for the attended versus the ignored speech stream during speech stimulus presentation (Fig. 3A).

In contrast, alpha power (8–12 Hz) lateralization reflected the side to which participants attended (Fig. 3B) (7–9, 29). Alpha power was relatively increased in the hemisphere on the same side as the locus of attention (i.e., ipsilateral) and relatively
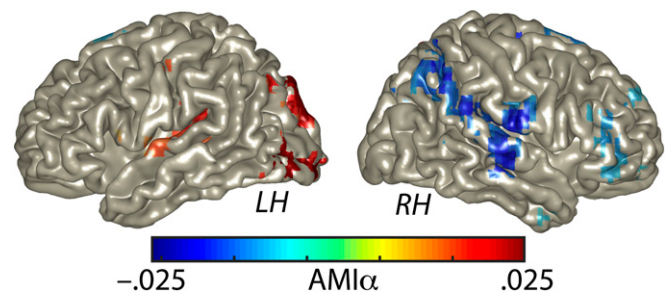
Fig. 4. Overlays on brain surfaces show the alpha power (8–12 Hz) AMI_α averaged across the time period of speech stimulus presentation (2.3–7.9 s). Warm colors indicate a relative increase in alpha power in attention-left compared with attention-right trials (and vice versa for cold colors). Overlays on the brain surfaces are masked at $P > 0.05$ (one-sample *t* tests of the attentional modulation index against zero; uncorrected). Note significant positive vs. negative AMI_α values in left vs. right auditory cortex regions, respectively.
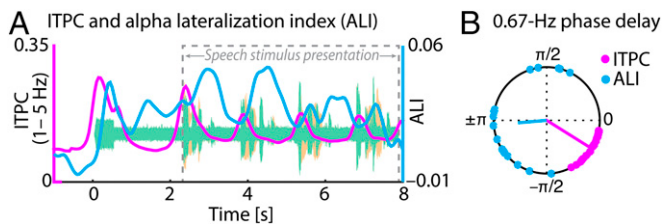
Fig. 5. (*A*) 1- to 5-Hz phase coherence (ITPC; averaged across 102 combined gradiometer sensors; magenta) and average ALI (cyan) superimposed on a stimulus waveform. Spoken digits were presented at a frequency of 0.67 Hz (i.e., ~1.49 s onset-to-onset delay of digits). (*B*) Phase angles of 0.67-Hz amplitude modulations of ITPC and ALI during speech stimulus presentation (2.3–7.9 s). Dots show 19 participants' phase angles; lines show mean phase angles, which differed significantly for ALI vs. ITPC (*P* < 0.001). An average 137.65° or 570-ms phase lag of ALI relative to ITPC was observed.

decreased in the hemisphere on the side that participants ignored (i.e., contralateral). Critically, the strength of alpha power lateralization varied over time and synchronized with the speech rate of the stimulus (Fig. 5). Thus, alpha power lateralization not only codes the locus of attention in space (10, 11, 13) but also the structure of the stimulus in time.

The differential temporal patterns of low-frequency phase and alpha power responses in our listening task discloses the neural processing of attended and ignored stimuli. Spatial attention did not affect phase-locked responses to speech, which suggests that to-be-attended but also ignored stimuli are initially encoded (30). Lateralization of alpha power speaks to the inhibition of neural activity related to ignored stimuli, mediated by high alpha power in the hemisphere ipsilateral to the locus of attention (and vice versa for low alpha power in the contralateral hemisphere) (24). The peaks of alpha power lateralization reveal that attention exerts the maximum impact on the neural activity ~0.6 s after the encoding-related phase coherence at digit onsets. The seemingly long time interval between stimulus onset and the attention-driven alpha lateralization is in agreement with the generally slowly evolving auditory alpha response from the literature (6, 31), which temporally lags behind the onsets of speech items by several hundreds of milliseconds (e.g., 32, 33). Furthermore, the high spectro-temporal overlap of simultaneously presented digits (same voice, same perceptual onset time) probably prolonged the perceptual segregation before selection of attended speech could commence (34). Our results show that alpha power lateralization implements a spatially specific filter that acts on neurally encoded competing sensory information to enhance neural activity related to attended stimuli.

**Attention to Speech Modulates Alpha Power in Auditory Cortex.** Spatial attention modulates alpha power in primary sensory cortices across modalities (8–10), but the cortical generators of alpha power modulation for spatial attention to speech have thus far not been identified. We found that spatial cues on one ear enhance low-frequency phase coherence in the contralateral hemisphere (Fig. 3*A*, cue period), which agrees with research showing that the left auditory cortex receives input primarily from the right ear and vice versa for right auditory cortex (35, 36). Critically, spatial attention to speech modulated alpha power in opposing ways in left and right auditory cortices (Fig. 4): Alpha power was relatively enhanced in the auditory cortex receiving its major input from the ear presented with the to-be-ignored speech (and vice versa for the other auditory cortex). Attention thus relatively enhances auditory processing of attended speech compared with ignored speech in auditory cortex (5). These results extend the thus far sparse evidence for a cognitive function of auditory alpha oscillations (18, 37, 38) by showing that they are controlled by top-down attention to speech.

As expected in a spatial attention task, we also found enhanced alpha power in the right parietal cortex when attention was directed to the right side compared with the left side. The parietal cortex codes the spatial location of stimuli in our environment (39, 40). Attention-driven alpha power lateralization speaks for an auditory attention network including the supramodal parietal cortex and auditory-specific temporal regions (41). However, our analysis of spatial attention on neural responses reveals only brain areas responding in opposing ways in "attention-left" versus "attention-right" trials. It is thus likely that the auditory attention network comprises nonlateralized top-down control regions in the frontal cortex (such as the frontal eye field) (42, 43) that drive the present parietal and auditory alpha power effects.

Spatial Alpha Power Modulations over Time Enhance Listening Performance. Task performance is known to relate to changes in alpha power during stimulus presentation (e.g., 6, 11). However, there is at present no evidence that temporal synchronization of alpha power with human speech supports the recall of attended speech items. Our results show that larger modulation of alpha power lateralization in synchrony with the stimulus predicts better recall of attended speech stimuli (Fig. 6). Furthermore, participants with a larger modulation of alpha power lateralization in incorrect trials (i.e., trials with at least one error) made fewer errors in these incorrect trials (i.e., recognized more of the four digits).

The behavioral benefit of enhanced alpha power modulation in synchrony with the stimulus implies that rhythmic rather than steady alpha power lateralization accompanies successful selection of attended stimuli (44). In general, a rhythmic oscillation between states of strong and weak modulation of neural activity is metabolically cheaper than strong modulation throughout time (45). Specifically, rhythmic lateralization of alpha power can implement rhythmic attentional sampling, as has been observed recently in the visual modality (46) for lateralized gamma activity (47). In line with the idea of temporally dynamic auditory attention (48), our results show that a neural signature of spatial attention (i.e., alpha power lateralization) synchronizes with the temporal structure of competing sensory information. Thereby, alpha power lateralization relatively enhances, in a pulsed manner, neural activity related to the attended stimuli in the time windows following sensory encoding.
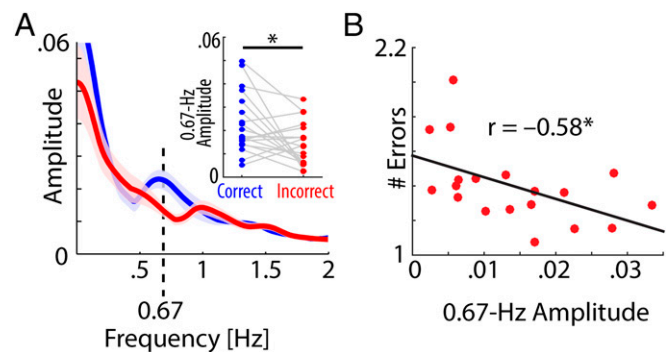
Fig. 6. (*A*) Amplitude spectra of the ALI during speech stimulus presentation (2.3–7.9 s) for correct (blue) and incorrect trials (red). Shaded areas show ±1 SEM. At the digit presentation rate of 0.67 Hz, the spectral amplitude was significantly enhanced for correct compared with incorrect trials (*Inset*; *P* = 0.017). (*B*) For incorrect trials, the amplitude of 0.67-Hz modulation of the ALI during speech stimulus presentation (2.3–7.9 s) was predictive of participants' average number of errors (*n* = 19; *P* = 0.01).

Conclusions. In sum, attention induces a spatially specific increase and decrease of alpha power in the respective hemisphere ipsilateral and contralateral to the locus of attended speech. Critically, this alpha power lateralization is also temporally specific because it is modulated in synchrony with the speech rate of the stimulus. Low-frequency phase coherence indicates that sensory encoding of attended and ignored stimuli is unselective but is followed in time by the spatio-temporally specific alpha power lateralization. Thus, alpha power modulations selectively enhance neural activity related to the attended stimulus in an attention network comprising supramodal parietal as well as sensory-specific regions (i.e., auditory cortex).

Materials and Methods

Participants. Nineteen young (median = 27 y; range = 23–34 y; 10 females) right-handed German native speakers took part in this study. Data of one additional participant were excluded from all analyses due to technical problems during MEG recording. Participants gave written informed consent and were financially compensated for their participation. Procedures were approved by the local ethics committee of the medical faculty of the University of Leipzig.

Stimuli. The spatial cue at the beginning of each trial was a 1,000-Hz pure tone of 500-ms duration (50-ms linear onset ramp) equal in intensity to the spoken digits. Recordings of German digits ranged between 21 and 99 (excluding integer multiples of 10; sampling rate: 44.1 kHz) and were adopted from previous studies (6, 49). Digits contained four syllables and had an average duration of 1.125 s. Root mean square intensity was equalized across digits. The perceptual center (P-center) (22) of each digit was defined as the time point when the digit signal's broad-band envelope (15-Hz lowpass-filtered modulus of the Hilbert transform) reached 50% of the first syllable's peak amplitude. The P-center was used to align words within two speech streams (see next section).

Procedure. During a trial, participants fixated a centrally presented cross. Trials started with a spatial cue tone on one side (left or right ear), indicating the to-be-attended side; 2.3 s after cue onset, four spoken digits were presented to the left and four different concurrent digits to the right side (Fig. 1A). For concurrent digits, perceptual centers were temporally aligned and digits were distinct in their 10 and one positions (e.g., co-occurences of "35" and "37" or "81" and "21" were avoided). The onset-to-onset interval of the digits was 1.49 s, resulting in a digit presentation rate of 0.67 Hz (50). Moderately quiet, continuous white noise (signal-to-noise ratio = +10 dB) masked cue and digits. Approximately 1 s (jittered 0.8–1.2 s) after the offset of the last digit a response screen was visually presented that contained 12 digits (four from the to-be-attended side, four from the to-be-ignored side, and four random digits not presented in any stream). Participants were asked to use a MEG-compatible trackball mouse (Logitech Marble Mouse) to select the four digits that appeared on the to-be-attended side in any order. Digits on the response screen were presented in either ascending or descending order (randomized from trial to trial) to prevent motor preparation during a trial. After selection of four digits, the next trial started self-paced after an additional mouse click. Auditory materials were presented via plastic ear tubes at an average intensity of ~70 dB sound pressure level. Visual stimuli were shown on a back projection screen. Each participant performed 150 trials (five blocks of 30 trials each). Trial order was fully randomized with the constraint that the spatial cue appeared on the left side in half of the trials within each block. Including preparation time, the experiment lasted approximately 2 h.

Behavioral Data Analyses. Due to technical reasons, no behavioral responses were registered for, on average, 1.2% (SD = 2.7%) of the trials per participant, which were removed from all further behavioral and MEG data analyses.

MEG Data Recording and Analyses. Participants were seated in a magnetically shielded room (Vaccumschmelze). A 306-sensor Neuromag Vectorview MEG (Elekta) measured magnetic fields at 102 locations from 204 orthogonal planar gradiometers and 102 magnetometers. Additionally, the electroencephalogram from 64 scalp electrodes (Ag/Ag-Cl) was recorded but not analyzed in this study. Each participant's head position was monitored with five head position indicator coils. MEG responses were recorded at a sampling rate of 1,000 Hz with a DC–300-Hz bandwidth. Offline, the signal space separation method (51) was used to suppress external disturbances (i.e., noise), interpolate bad sensors, and transform individual participant data to a common sensor space across blocks.

For subsequent MEG data analyses, we used the Fieldtrip toolbox (52) for Matlab (R2013b) and customized Matlab scripts. Only data recorded from gradiometer sensors were analyzed. Continuous data were highpass-filtered at 0.3 Hz (finite impulse response filter, FIR; zero-phase lag; order 5574; hanning window) and lowpass-filtered at 180 Hz (FIR; zero-phase lag; order 100; hamming window). Data were down-sampled to 500 Hz and epochs from −2 to 10 s around cue onset were extracted. Epochs were rejected when MEG responses at any gradiometer sensor exceeded 800 pT/m. Independent component analysis was used to reject components corresponding to eye blinks, saccadic eye movements, muscle activity, heartbeats, drifts, and jumps. Single-trial time series were convolved with a family of Morlet wavelets (1–20 Hz; in steps of 0.5 Hz and 0.05 s). To calculate ITPC (Fig. 2A), the complex wavelet coefficients (obtained with a wavelet width of three cycles) were divided by their magnitudes and averaged across trials, followed by calculating the magnitude of the resulting complex value (ITPC; Fig. 2A; 53). Time-frequency power representations (obtained with a wavelet width of seven cycles) were obtained by squaring the magnitude of the estimated complex wavelet transform coefficients. Data from 204 gradiometer sensors (102 pairs of gradiometer sensors) were combined by averaging ITPC and by summing power estimates for the two sensors at the same location.

Attention Indices. Two indices of attentional modulation of neural responses were calculated for each time-frequency bin (−2 to 10 s; 1–20 Hz). First, the attentional modulation index $[AMI_X = (X_{att\_left} − X_{att\_right})/(X_{att\_left} + X_{att\_right})]$ revealed a spatially resolved measure of attention effects on intertrial phase coherence (1–5 Hz; AMI_{ITPC}) and alpha power (8–12 Hz; AMI_α) at each sensor (X being either ITPC or α power; Fig. 3). Because spatial topographies of AMI_α differed across individuals, we selected 20 sensors with the largest positive AMI_α values on the left hemisphere and 20 sensors with the largest negative AMI_α values on the right hemisphere for each participant (8) (Fig. S4). For attention-left trials, left-hemispheric selected sensors were considered ipsilateral and right-hemispheric selected sensors were considered contralateral and vice versa for attention-right trials. The ipsi- and contralateral sensors were used to calculate the second attention index, that is, the alpha lateralization index $[ALI_\alpha = (\alpha_{att\_ipsi} − \alpha_{att\_contra})/(\alpha_{att\_ipsi} + \alpha_{att\_contra})]$. The ALI was calculated over the time course of the trial using a 250-ms sliding window (54) and revealed a temporally resolved measure of attention effects on alpha power (Fig. 5A). Note that both indices were not calculated for single trials, but for each participant's mean 8- to 12-Hz power across attention-left/attention-right trials (for AMI_α) and mean 1- to 5-Hz ITPC at ipsilateral/contralateral sensors (for AMI_{ITPC}).

Source Analyses. Individual cortical surfaces and inner skull surfaces were created from each participant's T1-weighted MRI scans (using Freesurfer and MNE software). The MR and the MEG coordinate systems were coregistered using more than 60 digitized points on the head surface. The brain volume of each participant was normalized to a common standardized source model and divided into a grid of 4-mm resolution. Leadfields were calculated for all grid points (55). We applied the Dynamic Imaging of Coherent Sources beamformer approach (27) implemented in the Fieldtrip toolbox. For the source analysis of overall alpha power (Fig. 2D), we calculated the cross-spectral density of Fast Fourier Transforms at 10 Hz with ±2 Hz spectral smoothing (time period: 0–7.9 s). Using each participant's leadfield and cross-spectral density, a spatially adaptive filter was calculated for each of the 37,163 source locations inside the brain arranged in a regular grid of 4-mm steps. This filter was applied to Fourier-transformed data to estimate alpha power for each grid point. The neural activity index was calculated by dividing the activation at each source location by its noise estimate (56). For source analysis of ITPC (Fig. 2B), we derived a spatially adaptive filter based on each participant's leadfield and the cross-spectral density of Fast Fourier Transforms centered at 3 Hz with ±2 Hz spectral smoothing (time period: 0–7.9 s). This filter was applied to single-trial Fourier Transforms (1–5 Hz, in steps of ~0.13 Hz). ITPC at each grid point was calculated and averaged across frequencies.

For localization of the AMI_α, we calculated the cross-spectral density of Fast Fourier Transforms at 10 Hz with ±2 Hz spectral smoothing separately for attention-left and attention-right trials for three time periods (cue: 0–0.5 s; anticipation: 0.5–2.3 s; speech stimulus presentation: 2.3–7.9 s). For each time period, a common spatial filter was calculated on the basis of trials under all conditions (attention-left and -right) for each participant, using the leadfield and cross-spectral density. The common filter of a specific time period was used for source projection of attention-left and attention-right conditions separately. AMI_α was calculated for each participant ($n = 19$) at each grid point: $[AMI_\alpha = (\alpha_{att\_left} − \alpha_{att\_right})/(\alpha_{att\_left} + \alpha_{att\_right})]$ and tested for significant difference from zero (one-sample t test; for visualization, AMI_α values with $P > 0.05$ were masked) (Fig. 4 and Fig. S3).

NEUROSCIENCE

Spectral Analysis of ITPC and ALI. Fast Fourier Transforms were calculated for ITPC and ALI time courses during speech stimulus presentation (2.3–7.9 s, using zero-padding to obtain a frequency resolution of 0.01 Hz) (Fig. 5). The magnitude of the resulting complex coefficients was used as measure of frequency-specific modulation strength. For the ITPC analysis of correct and incorrect trials (Fig. S9), ITPC time courses, Fast Fourier Transforms, and frequency-specific modulation, strengths were calculated for 1,000 random draws of the maximum equal number of correct and incorrect trials of each participant, followed by averaging across the 1,000 draws. This procedure controlled for the effect of trial number on ITPC estimates.

Statistical Testing. We applied parametric tests when the data conformed to normality assumptions ($P > 0.05$ in Shapiro–Wilk test) and appropriate nonparametric tests otherwise. For effect sizes, we report the $r_{equivalent}$ (bound between 0 and 1) (57), which is equal to the Pearson product-moment correlation for two continuous variables, to the point-biserial correlation for one continuous and one dichotomous variable, and to Kendall's coefficient of concordance for the nonparametric Friedman test. For circular statistics, we report the circular–linear correlation of phase angles and condition labels. For circular Rayleigh tests, we report the resultant vector length (r).

1. Ward LM (2003) Synchronous neural oscillations and cognitive processes. *Trends Cogn Sci* 7(12):553–559.
2. Lakatos P, et al. (2013) The spectrotemporal filter mechanism of auditory selective attention. *Neuron* 77(4):750–761.
3. Fries P, Reynolds JH, Rorie AE, Desimone R (2001) Modulation of oscillatory neuronal synchronization by selective visual attention. *Science* 291(5508):1560–1563.
4. Foxe JJ, Simpson GV, Ahlfors SP (1998) Parieto-occipital approximately 10 Hz activity reflects anticipatory state of visual attention mechanisms. *Neuroreport* 9(17):3929–3933.
5. Strauss A, Wöstmann M, Obleser J (2014) Cortical alpha oscillations as a tool for auditory selective inhibition. *Front Hum Neurosci* 8:350.
6. Wöstmann M, Herrmann B, Wilsch A, Obleser J (2015) Neural alpha dynamics in younger and older listeners reflect acoustic challenges and predictive benefits. *J Neurosci* 35(4):1458–1467.
7. Worden MS, Foxe JJ, Wang N, Simpson GV (2000) Anticipatory biasing of visuospatial attention indexed by retinotopically specific alpha-band electroencephalography increases over occipital cortex. *J Neurosci* 20(6):RC63.
8. Haegens S, Händel BF, Jensen O (2011) Top-down controlled alpha band activity in somatosensory areas determines behavioral performance in a discrimination task. *J Neurosci* 31(14):5197–5204.
9. Bauer M, Kennett S, Driver J (2012) Attentional selection of location and modality in vision and touch modulates low-frequency activity in associated sensory cortices. *J Neurophysiol* 107(9):2342–2351.
10. Frey JN, et al. (2014) Selective modulation of auditory cortical alpha activity in an audiovisual spatial attention task. *J Neurosci* 34(19):6634–6639.
11. Kerlin JR, Shahin AJ, Miller LM (2010) Attentional gain control of ongoing cortical speech representations in a "cocktail party." *J Neurosci* 30(2):620–628.
12. Romei V, Gross J, Thut G (2010) On the role of prestimulus alpha rhythms over occipito-parietal areas in visual input regulation: Correlation or causation? *J Neurosci* 30(25):8692–8697.
13. Thut G, Nietzel A, Brandt SA, Pascual-Leone A (2006) Alpha-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *J Neurosci* 26(37):9494–9502.
14. Herrmann B, Henry MJ, Grigutsch M, Obleser J (2013) Oscillatory phase dynamics in neural entrainment underpin illusory percepts of time. *J Neurosci* 33(40):15799–15809.
15. Henry MJ, Obleser J (2012) Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. *Proc Natl Acad Sci USA* 109(49):20095–20100.
16. Luo H, Poeppel D (2007) Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron* 54(6):1001–1010.
17. Ahissar E, et al. (2001) Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proc Natl Acad Sci USA* 98(23):13367–13372.
18. Herrmann B, Henry MJ, Haegens S, Obleser J (2016) Temporal expectations and neural amplitude fluctuations in auditory cortex interactively influence perception. *NeuroImage* 124(Pt A):487–497.
19. Ding N, Simon JZ (2012) Emergence of neural encoding of auditory objects while listening to competing speakers. *Proc Natl Acad Sci USA* 109(29):11854–11859.
20. Zion Golumbic EM, et al. (2013) Mechanisms underlying selective neuronal tracking of attended speech at a "cocktail party." *Neuron* 77(5):980–991.
21. Arnal LH, Wyart V, Giraud AL (2011) Transitions in neural oscillations reflect prediction errors generated in audiovisual speech. *Nat Neurosci* 14(6):797–801.
22. Morton J, Marcus S, Frankish C (1976) Perceptual centers (P-centers). *Psychol Rev* 83(5):405–408.
23. Fu KM, et al. (2001) 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(1):145–152.
24. Jensen O, Mazaheri A (2010) Shaping functional architecture by oscillatory alpha activity: Gating by inhibition. *Front Hum Neurosci* 4:186.
25. Mazaheri A, et al. (2014) Region-specific modulations in oscillatory alpha activity serve to facilitate processing in the visual and auditory modalities. *Neuroimage* 87:356–362.
26. Adrian ED (1944) Brain rhythms. *Nature* 153:360–362.
27. Gross J, et al. (2001) Dynamic imaging of coherent sources: Studying neural interactions in the human brain. *Proc Natl Acad Sci USA* 98(2):694–699.
28. Ng BS, Logothetis NK, Kayser C (2013) EEG phase patterns reflect the selectivity of neural firing. *Cereb Cortex* 23(2):389–398.
29. Ahveninen J, Huang S, Belliveau JW, Chang WT, Hämäläinen M (2013) Dynamic oscillatory processes governing cued orienting and allocation of auditory attention. *J Cogn Neurosci* 25(11):1926–1943.
30. Lavie N (2005) Distracted and confused?: Selective attention under load. *Trends Cogn Sci* 9(2):75–82.
31. Shahin AJ, Picton TW, Miller LM (2009) Brain oscillations during semantic evaluation of speech. *Brain Cogn* 70(3):259–266.
32. Obleser J, Weisz N (2012) Suppressed alpha oscillations predict intelligibility of speech and its acoustic details. *Cereb Cortex* 22(11):2466–2477.
33. Krause CM (2006) Cognition- and memory-related ERD/ERS responses in the auditory stimulus modality. *Prog Brain Res* 159:197–207.
34. Shinn-Cunningham BG (2008) Object-based auditory and visual attention. *Trends Cogn Sci* 12(5):182–186.
35. Rosenzweig MR (1951) Representations of the two ears at the auditory cortex. *Am J Physiol* 167(1):147–158.
36. Tervaniemi M, Hugdahl K (2003) Lateralization of auditory-cortex functions. *Brain Res Brain Res Rev* 43(3):231–246.
37. Weisz N, Hartmann T, Müller N, Lorenz I, Obleser J (2011) Alpha rhythms in audition: Cognitive and clinical perspectives. *Front Psychol* 2:73.
38. Weisz N, Müller N, Jatzev S, Bertrand O (2014) Oscillatory alpha modulations in right auditory regions reflect the validity of acoustic cues in an auditory spatial attention task. *Cereb Cortex* 24(10):2579–2590.
39. Colby CL, Goldberg ME (1999) Space and attention in parietal cortex. *Annu Rev Neurosci* 22:319–349.
40. Vallar G (1998) Spatial hemineglect in humans. *Trends Cogn Sci* 2(3):87–97.
41. Banerjee S, Snyder AC, Molholm S, Foxe JJ (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(27):9923–9932.
42. Müller N, Weisz N (2012) Lateralized auditory cortical alpha band activity and interregional connectivity pattern reflect anticipation of target sounds. *Cereb Cortex* 22(7):1604–1613.
43. Marshall TR, O'Shea J, Jensen O, Bergmann TO (2015) Frontal eye fields control attentional modulation of alpha and gamma oscillations in contralateral occipitoparietal cortex. *J Neurosci* 35(4):1638–1647.
44. Merchant H, Grahn J, Trainor L, Rohrmeier M, Fitch WT (2015) Finding the beat: A neural perspective across humans and non-human primates. *Philos Trans R Soc Lond B Biol Sci* 370(1664):20140093.
45. Buzsáki G, Draguhn A (2004) Neuronal oscillations in cortical networks. *Science* 304(5679):1926–1929.
46. Landau AN, Fries P (2012) Attention samples stimuli rhythmically. *Curr Biol* 22(11):1000–1004.
47. Landau AN, Schreyer HM, van Pelt S, Fries P (2015) Distributed attention is implemented through theta-rhythmic gamma modulation. *Curr Biol* 25(17):2332–2337.
48. Large EW, Jones MR (1999) The dynamics of attending: How people track time-varying events. *Psychol Rev* 106(1):119–159.
49. Wöstmann M, Schröger E, Obleser J (2015) Acoustic detail guides attention allocation in a selective listening task. *J Cogn Neurosci* 27(5):988–1000.
50. Lakatos P, Schroeder CE, Leitman DI, Javitt DC (2013) Predictive suppression of cortical excitability and its deficit in schizophrenia. *J Neurosci* 33(28):11692–11702.
51. Taulu S, Kajola M, Simola J (2004) Suppression of interference and artifacts by the Signal Space Separation Method. *Brain Topogr* 16(4):269–275.
52. Oostenveld R, Fries P, Maris E, Schoffelen JM (2011) FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci* 2011:156869.
53. Lachaux JP, Rodriguez E, Martinerie J, Varela FJ (1999) Measuring phase synchrony in brain signals. *Hum Brain Mapp* 8(4):194–208.
54. Mesgarani N, Chang EF (2012) Selective cortical representation of attended speaker in multi-talker speech perception. *Nature* 485(7397):233–236.
55. Nolte G (2003) The magnetic lead field theorem in the quasi-static approximation and its use for magnetoencephalography forward calculation in realistic volume conductors. *Phys Med Biol* 48(22):3637–3652.
56. Van Veen BD, van Drongelen W, Yuchtman M, Suzuki A (1997) Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Trans Biomed Eng* 44(9):867–880.
57. Rosenthal R, Rubin DB (2003) r equivalent: A simple effect size indicator. *Psychol Methods* 8(4):492–496.
58. Kimura D (1961) Cerebral dominance and the perception of verbal stimuli. *Can J Psychol* 15(3):166–171.

Supporting Information

Wöstmann et al. 10.1073/pnas.1523357113

Further Analysis of Behavioral Performance

To compare performance in attention-left vs. attention-right trials, a nonparametric Friedman test was applied to the difference between attention conditions (left–right) for the three response types (hit, spatial confusion, random error), which was significant (χ^2 (2) = 24; $P < 0.001$; $r = 0.63$). Post hoc Wilcoxon signed-rank tests were used to compare proportions of hits, spatial confusions, and random errors between attention-left and attention-right trials (Fig. S1). The proportion of hits was higher in attention-right compared with attention-left trials ($z = 3.7$; $P < 0.001$; $r = 0.24$). However, proportions of spatial confusions ($z = 3.67$; $P < 0.001$; $r = 0.31$) and random errors ($z = 2.07$; $P = 0.039$; $r = 0.13$) were higher in attention-left trials. These results demonstrate the right ear advantage; that is, recall is enhanced for speech materials presented to the right compared with left ear during dichotic listening (58).
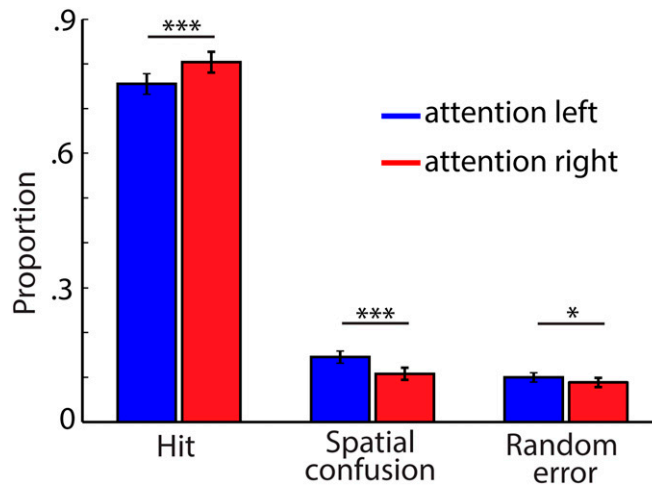
Fig. S1. Bars show average proportions of three response types (hit, spatial confusion, random error), separately for attention-left (blue) and attention-right trials (red). Error bars indicate ±1 SEM ***$P < 0.001$; *$P < 0.05$.
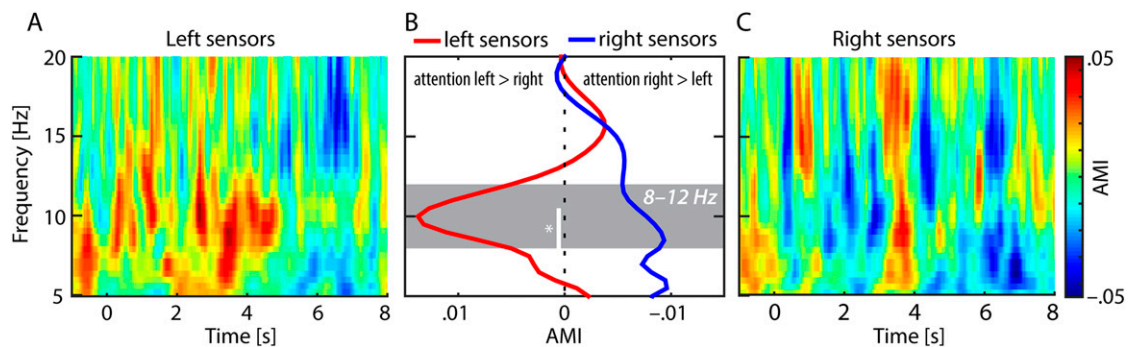

Fig. S2. AMI for frequencies 5–20 Hz averaged across all sensors (*A*) on the left hemisphere and (*C*) on the right hemisphere. (*B*) Mean AMI across time of the dichotic listening task (cue onset, 0 s; last digit offset, 7.9 s), for left sensors (red) and right sensors (blue). The white bar indicates frequencies for which the AMI differed significantly between left and right sensors (8–10.5 Hz; $P < 0.05$; paired *t* tests, FDR-corrected for multiple comparisons). The alpha frequency band (8–12 Hz) is highlighted in gray.
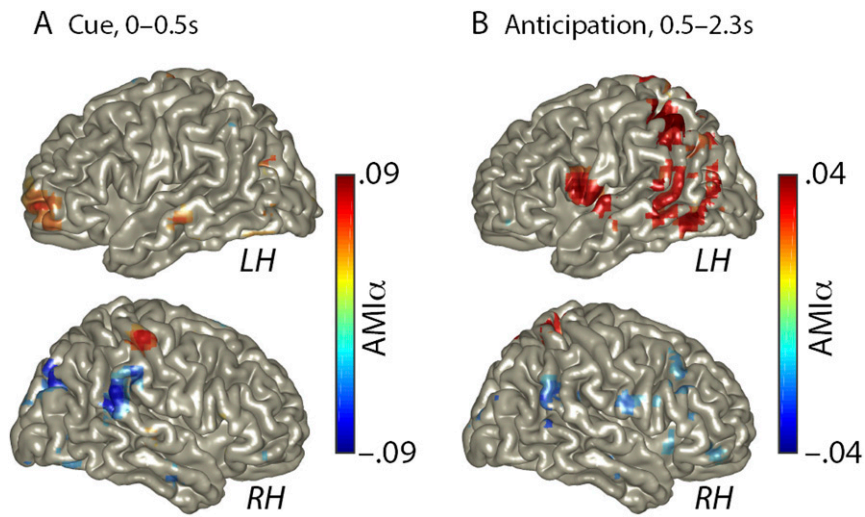
Fig. S3. Overlays on brain surfaces show the alpha power (8–12 Hz) AMI_α averaged across the time period of (*A*) cue presentation (0–0.5 s) and (*B*) stimulus anticipation (0.5–2.3 s). Warm colors indicate a relative increase in alpha power in attention-left compared with attention-right trials (and vice versa for cold colors). Overlays on the brain surfaces are masked at $P > 0.05$ (one-sample t tests of the AMI_α against zero; uncorrected).
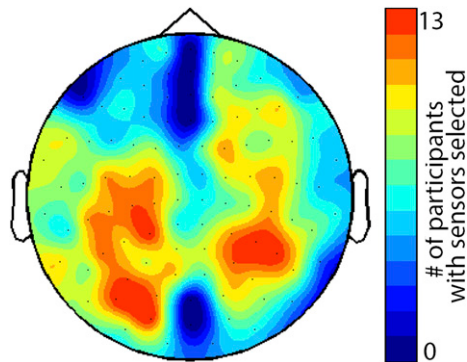
Fig. S4. To calculate the ALI, 20 left-hemispheric and 20 right-hemispheric sensors were selected for each individual (for details, see main text). The topographic map shows the number of subjects for whom individual sensors were selected (total number of subjects, $n = 19$). Note that centro-parietal sensors on both hemispheres were selected most frequently.

A Single subject AMIα & sensor selection
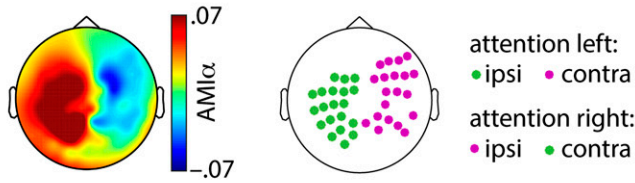
$$\text{Alpha lateralization index (ALI)} = \frac{\alpha_{att\_ipsi} - \alpha_{att\_contra}}{\alpha_{att\_ipsi} + \alpha_{att\_contra}}$$

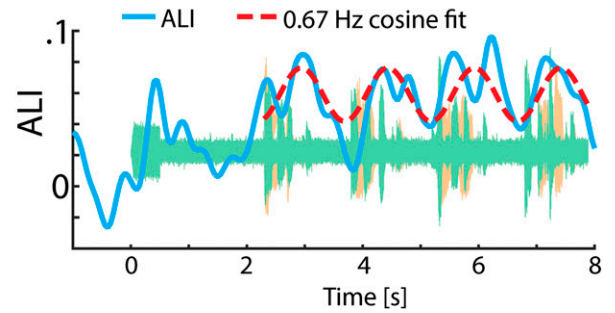## B Single subject Alpha lateralization index

Fig. S5. (*A, Left*) Topographic map of the average alpha power (8–12 Hz) AMI$_\alpha$ across time of the dichotic listening task (0–7.9 s) for a single participant. (*A, Right*) Twenty sensors were selected on the left and right hemisphere that exhibited maximal and minimal AMI$_\alpha$ values, respectively. Green and purple dots show selected sensors on the left and right hemisphere, respectively. Depending on the condition (attention left vs. attention right), these sensors were classified as ipsi (-lateral) and contra (-lateral) to calculate the ALI. (*B*) ALI for the subject shown in *A* superimposed on a stimulus waveform. The alpha lateralization index (cyan) exhibited characteristic modulations over the trial time course. During the speech stimulus presentation (2.3–7.9 s), the alpha lateralization index was modulated at the digit presentation rate (0.67 Hz), highlighted for illustration purposes by a 0.67-Hz cosine fit (dashed red line).
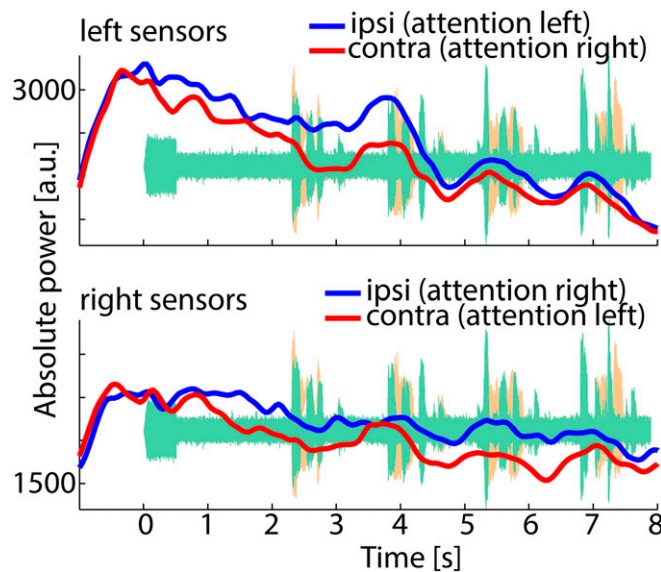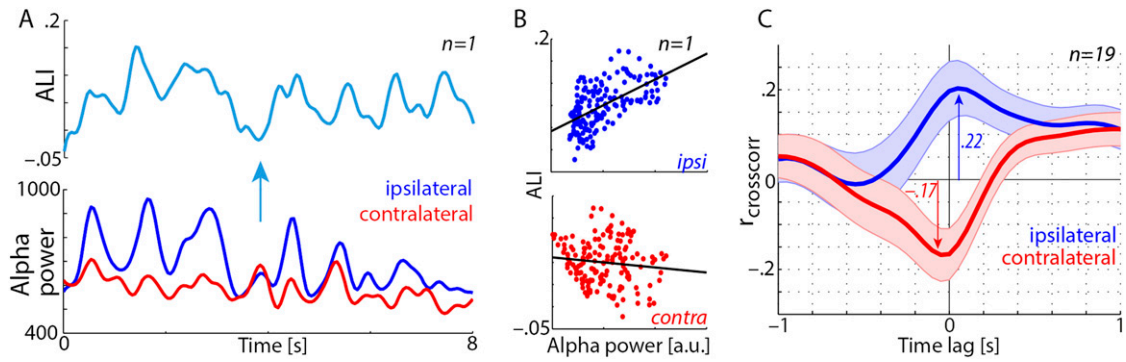

Fig. S6. Time courses of alpha power (8–12 Hz) at ipsilateral (blue) and contralateral (red) sensors superimposed on a stimulus waveform. Note that participant-individual sensors (shown in Fig. S4 and Fig. 5A) were used for the calculation of ipsi- and contralateral alpha power time courses.

Fig. S7. Temporal modulation of the ALI is driven to equal extents by alpha power changes at sensors ipsilateral and contralateral to the attended side. (*A*) ALI (cyan, *Top*) and alpha power at ipsilateral (blue) and contralateral sensors (red, *Bottom*) for one representative subject. (*B*) Scatterplots show the direct relationship of ipsi- (*Top*) and contralateral (*Bottom*) alpha power to the ALI, respectively (for 50-ms time steps from cue onset, 0 s, until trial end, 8 s), from the same subject as *A*. Least-squares regression lines (black) demonstrate a positive relationship between the ALI and ipsilateral alpha power (*Top*) and a negative relationship between the ALI and contralateral alpha power (*Bottom*). (*C*) To explore potential time-lagged correlations between ALI and ipsi-/contralateral alpha power, solid lines and shaded areas show the cross-correlation (mean ±1 SEM) between ALI and the ipsi- (blue) and contralateral (red) alpha power, respectively, averaged across all $n = 19$ subjects. Correlation coefficients of $r_{crosscorr} = 0$ would indicate no linear relationship between ipsi/contralateral alpha power and the ALI. Positive time lags indicate a temporal delay of the ALI with respect to ipsi/contralateral alpha power. Critically, for time lags close to 0 s (i.e., negligible temporal delay of the ALI relative to alpha power), ipsilateral alpha power correlates positively with the ALI (one sample t test of Fisher z-transformed $r_{crosscorr}$ values at time lag 0 s; $t_{18} = 2.9$; $P = 0.01$; $r = 0.44$; strongest positive correlation, $r_{crosscorr} = 0.22$), whereas contralateral alpha power correlates with equal magnitude negatively with the ALI ($t_{18} = 2.8$; $P = 0.01$; $r = 0.42$; strongest negative correlation, $r_{crosscorr} = -0.17$). These results demonstrate that the positive peaks of the ALI (such as shown in Fig. 5) are driven by decreasing contralateral but also increasing ipsilateral alpha power.
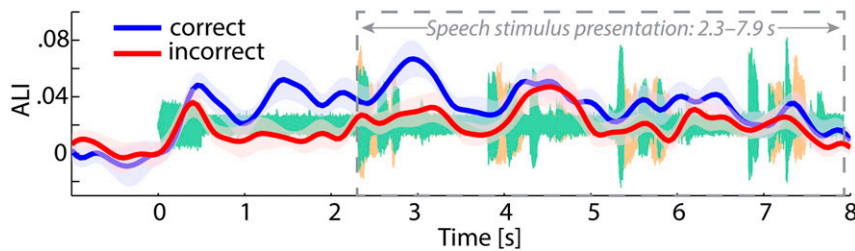
Fig. S8. Alpha lateralization index for correct (blue) and incorrect trials (red) superimposed on a stimulus waveform. Shaded areas indicate ±1 SEM.
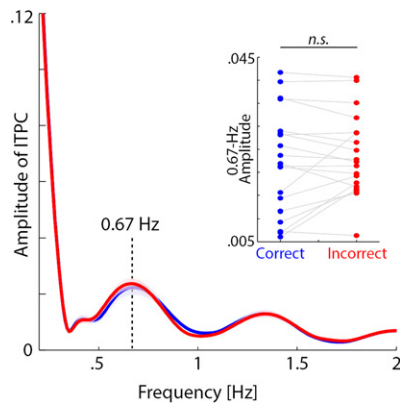
Fig. S9. Fourier spectra of 1- to 5-Hz ITPC during selective listening (2.3–7.9 s) for correct (blue) and incorrect trials (red) in analogy to the respective results for the alpha lateralization index (Fig. 6). Because trial number affects ITPC, Fourier spectra were estimated from the average of spectra calculated for 1,000 random draws with an equal number of correct and incorrect trials for each participant. Shaded areas indicate ±1 SEM. Spectral amplitude of 1–5 Hz ITPC at the 0.67-Hz digit presentation rate did not differ between correct and incorrect trials (*Inset*; paired-samples t test, $t_{18} = 1.12$; $P = 0.276$; $r = 0.07$), and no significant 0.67-Hz phase delay between 1- to 5-Hz ITPC for correct and incorrect trials was observed (Parametric Hotelling paired-sample test; $F_{34} = 1.12$; $P = 0.348$; $r = 0.23$).