Suppression of irrelevant sounds during auditory working memory

Jyrki Ahveninen[1], **Larry J. Seidman**[2,3], **Wei-Tang Chang**[1], **Matti Hämäläinen**[1,4], and **Samantha Huang**[1]

[1]Harvard Medical School – Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital, Charlestown, MA

[2]Massachusetts Mental Health Center Public Psychiatry Division, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA

[3]Department of Psychiatry, Harvard Medical School, Massachusetts General Hospital, Boston, MA

[4]Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, MA, USA

Abstract

Auditory working memory (WM) processing in everyday acoustic environments depends on our ability to maintain relevant information online in our minds, and to suppress interference caused by competing incoming stimuli. A challenge in communication settings is that the relevant content and irrelevant inputs may emanate from a common source, such as a talkative conversationalist. An open question is how the WM system deals with such interference. Will the distracters become inadvertently filtered before processing for meaning because the primary WM operations deplete all available processing resources? Or are they suppressed post perceptually, through an active control process? We tested these alternative hypotheses by measuring magnetoencephalography (MEG), EEG, and functional MRI (fMRI) during a phonetic auditory continuous performance task. Contextual WM maintenance load was manipulated by adjusting the number of "filler" letter sounds in-between cue and target letter sounds. Trial-to-trial variability of pre- and post-stimulus activations in fMRI-informed cortical MEG/EEG estimates was analyzed within and across 14 subjects using generalized linear mixed effect (GLME) models. High contextual WM maintenance load suppressed left auditory cortex (AC) activations around 250–300 ms after the onset of irrelevant phonetic sounds. This effect coincided with increased 10–14 Hz alpha-range oscillatory functional connectivity between the left dorsolateral prefrontal cortex (DLPFC) and left AC. Suppression of AC responses to irrelevant sounds during active maintenance of the task context also correlated with increased pre-stimulus 7–15 Hz alpha power. Our results suggest that under high auditory WM load, irrelevant sounds are suppressed through a "late" active suppression

Corresponding Author: Jyrki Ahveninen, Ph.D., MGH/MIT/HMS-Martinos Center, Bldg. 149 13th Street, Charlestown MA 02129, Phone (617) 726-6584; fax (617) 726-7422, jyrki@nmr.mgh.harvard.edu.

mechanism, which prevents short-term consolidation of irrelevant information without affecting the initial screening of potentially meaningful stimuli. The results also suggest that AC alpha oscillations play an inhibitory role during auditory WM processing.

Introduction

Our ability to suppress irrelevant sounds from interfering with auditory working memory (WM), while also maintaining the capacity to monitor the acoustic environment, is crucial in everyday communication settings. For example, during a conversation, a listener may hear an interesting detail that needs to be held online and protected from auditory interference until the talker has finished. Despite many previous WM studies involving auditory stimuli, several fundamental questions on how the human brain achieves these feats are still open.

It has been long debated whether irrelevant inputs are filtered before ("early selection"; Broadbent, 1958; Hillyard et al., 1973; Woldorff et al., 1993) or after processing for meaning ("late selection"; Deutsch and Deutsch, 1963; Näätänen, 1992). A potential resolution is offered by the "load theory" of visual attention (Lavie, 2005), which suggests that the selection stage depends on contextual factors. For example, conditions with multiple distractors, i.e., high "perceptual load" will inherently promote "early selection" because no processing resources are left for irrelevant stimuli. In the auditory domain, this idea is supported by studies of the auditory cortex (AC) change detection response, the mismatch negativity (MMN). During easy conditions such as movie viewing MMN is elicited even to semantic changes in the unattended stream, allowing late selection (Näätänen et al., 2001; Pulvermuller et al., 2001), but MMN to task-irrelevant deviants is virtually abolished during difficult dichotic listening (i.e., "early selection") (Woldorff et al., 1991; Woldorff et al., 1998). Interestingly, recent studies suggest that auditory distracters could be filtered at an early stage, not only under perceptual, but also under high visual WM load (Halin et al., 2015; SanMiguel et al., 2008). If true, these findings would also imply that suppression of irrelevant information is a "byproduct" of the primary WM maintenance and manipulation operations, requiring no dedicated control processes.

However, in many everyday settings, relevant and irrelevant sounds occur in the same stream, and all inputs may need to be initially screened to allow selective WM processing (*i.e.*, "late selection"). For example, in the above-mentioned social conversation example, it would not be constructive for the listener to completely shut off the other person, even though the talker's voice interferes with their auditory WM. Consistent with these notions, experiments that emulate conditions where relevant and irrelevant sounds are embedded in overlapping streams suggest prominent auditory ERP (Berti and Schroger, 2003; Muller-Gass and Schroger, 2007) and behavioral (Dalton et al., 2009) effects to irrelevant distracters also under high WM load. This has lead to an interpretation that distracter suppression requires WM resources and active control (Berti and Schroger, 2003; Dalton et al., 2009). According to visual studies, one of the areas controlling sensory gating specifically during WM maintenance could be the dorsolateral prefrontal cortex (DLPFC) (Postle, 2005).

Previous studies suggest that suppression of irrelevant information during attention and WM tasks is supported by neuronal oscillations at the alpha range (7–15 Hz) (Cooper et al., 2003;

Klimesch et al., 2007; Palva and Palva, 2007; Pfurtscheller, 2003). Alpha-band oscillations have been suggested to regulate cortical excitability to prioritize relevant and disregard irrelevant visual field locations (Foxe et al., 1998; Jensen and Mazaheri, 2010; Worden et al., 2000) and to suppress anticipated WM distracters (Bonnefond and Jensen, 2012). Similar effects are yet to be confirmed in the auditory domain. There is some evidence of lateralized alpha-power increases in ACs ipsilateral to the attended ear (Müller and Weisz, 2012). However, only a few previous studies have been able to show a direct relationship between AC alpha oscillations and distracter-elicited neuronal responses (Wöstmann et al., 2016). The alpha inhibition hypothesis is contrasted by an alternative interpretation that alpha synchronization enhances stimulus maintenance at the network level (Leiberg et al., 2006; Obleser et al., 2012; Palva and Palva, 2007; Wilsch and Obleser, 2015). This idea is based on correlations between alpha power and increased WM load, which has been observed both in visual (Palva and Palva, 2007) and auditory domains (Leiberg et al., 2006; Obleser et al., 2012; Wilsch and Obleser, 2015).

Here, we examined whether irrelevant sounds become inadvertently filtered before becoming fully processed during high auditory WM load ("filtering hypothesis"), or whether they are actively suppressed post-perceptually ("active suppression hypothesis"). In addition to examining the correlates of trial-to-trial variation during auditory WM, we also measured functional coupling between frontal and AC areas after filler sounds by analyzing long-range synchronization of neuronal oscillations, suggested to play a crucial role in WM (Jensen et al., 2007). Second, we also examined the role of alpha oscillations in suppression of irrelevant sensory information in ACs. Alternative hypotheses were compared using a multimodal approach, which combines information from MEG, EEG, and fMRI to achieve high-resolution estimates of stimulus processing and interregional connectivity in the cortical "source space".

In previous studies, the effects of WM load on auditory stimulus processing have often been studied using intermodal tasks (e.g., Halin et al., 2015; SanMiguel et al., 2008). However, in real-life settings, such as our initial conversation example, the distracters and relevant stimuli can be embedded into the same auditory stream that emanates from a common source. Here, to examine challenges under such conditions we used an auditory "AX" continuous performance task, modified from a paradigm sensitive to auditory WM changes on disorders such as schizophrenia (Seidman et al., 1998; Seidman et al., 2012; Seidman et al., 2016) (Fig. 1). During this task, the subjects monitor a sequence of spoken letters for a target letter, and respond to the target letter only when it follows a cue letter and pre-specified number of "filler" sounds (Seidman et al., 1998; Seidman et al., 2012; Seidman et al., 2016). This requires active maintenance of the cue and sequence information and suppression of the irrelevant fillers during the delay period. In other words, instead of "item load" in the classic sense, the current task requires active maintenance of the sequential task context. The longer the sequence during which this task context needs to be held online, the more it presumably burdens the WM system. In certain blocks the cue–target period could also be interleaved with additional cues, ensuring that the subject needed to continuously monitor the letter sounds also during the cue-target periods, instead of resorting to a simple counting strategy.

Methods

Task and design

Sixteen right-handed college educated adults with self-reported normal hearing and no neurological disorders, psychiatric conditions, or learning disabilities participated. Two subjects were excluded from the final sample due to failing to perform the task (hit rate <75% in a simple vigilance condition), rendering a total of 14 subjects for analysis (age 19–28 years, mean 21 years; 8 females). During an auditory continuous performance task (Seidman et al., 1998; Seidman et al., 2012) (Figure 1), the subjects monitored a sequence of voiced letter sounds for the presence of a target letter "A" that followed a cue letter "Q" and a pre-specified number of "filler" sounds that consisted of the other letters of the English alphabet. In vigilance control task blocks, the subjects were asked to respond to each "A" stimulus that occurred immediately after a "Q" event. In "low WM load" blocks, they were asked to respond to each "A" that occurred after a "Q" event and one other letter stimulus. Finally, in "high WM load" blocks, the subjects were to respond to each "A" stimulus that occurred after a "Q" event and three other letters. In this study, we were specifically interested how the maintenance of the auditory WM task set during this Q–A memory period modulates responses to task-irrelevant "filler" sounds. This interest is motivated by a recent multicenter study (Seidman et al., 2016), which suggested that poor performance in auditory "AX" tasks manipulating the WM maintenance load is among the best neuropsychological predictors of clinical psychosis in individuals at high risk for schizophrenia.

In addition to the number of filler sounds between the "Q" and "A", our factorial task also manipulated the level of higher-level control demands: In certain task blocks, the Q–A maintenance period was occasionally interrupted by distracter Q events or non-target A letters. In reference to our conversation example, this would be analogous to a significant event that occurs while the listener is holding an important thought and only cursorily monitoring the talker's voice. The task was divided to 90 s blocks during which the number of filler letters between the cue "Q" and target "A" remained the same. These task blocks started with a visual instruction cue. This instruction was identical before task blocks with or without interfering Q events. This ensured that the subject needed to keep track of the filler sounds in all task conditions.

The voiced letter sounds were presented at a comfortable level and at an average inter-stimulus interval (ISI) of 1 sec, jittered ±50 ms to prevent a buildup of subject's expectation. The original sound tokens, obtained from the Psychology Experiment Building Language (PEBL) Sound Archive version 0.1, had been edited to ensure the duration of each sound file was 400 ms. Similar task and stimuli were used in separate MEG/EEG and fMRI localizer sessions. The MEG/EEG sessions consisted of two 30-minute runs. The different tasks were presented as randomly ordered 90 s blocks, each starting with a 2-s visual instruction. The fMRI tasks were otherwise similar but the task instruction was changed after every third 10.5-s sparse sampling acquisition to avoid too lengthy analysis blocks in the fMRI analyses. It is important to note that, here, the fMRI data were used merely to confine the MEG/EEG source space, based on a pooled all vs. baseline contrast. In other words, the

comparisons of MEG/EEG source activities across conditions were independent of any differences in the corresponding fMRI effects.

Data acquisitions

Human subjects' approval was obtained and voluntary consents were signed before each measurement. 306-channel MEG (Elekta-Neuromag, Helsinki, Finland) and 74-channel EEG data were recorded simultaneously (600 samples/s, passband 0.01–192 Hz) in a magnetically shielded room. Common-average reference was utilized for all analyses of EEG data. The position of the head relative to the sensor array was monitored continuously using four Head-Position Indicator (HPI) coils attached to the scalp. Electro-oculogram (EOG) was also recorded to monitor eye artifacts. Whole-head 3T fMRI was acquired in a separate session using a 32-channel coil (Siemens TimTrio, Erlagen, Germany). To circumvent response contamination by scanner noise, we used a sparse-sampling gradient-echo BOLD sequence (TR/TE= 10,500/30 ms, 8.32 s silent period between acquisitions, flip angle 90°, FOV 192 mm) with 36 axial slices aligned along the anterior-posterior commissure line (3-mm slices, 0.75-mm gap, 3×3 mm^2 in-plane resolution), with the coolant pump switched off. T1-weighted anatomical images were obtained for combining anatomical and functional data using a multi-echo MPRAGE pulse sequence (TR=2510 ms; 4 echoes with TEs=1.64 ms, 3.5 ms, 5.36 ms, 7.22 ms; 176 sagittal slices with $1 \times 1 \times 1$ mm^3 voxels, 256×256 mm^2 matrix; flip angle = 7°). A field mapping sequence (TR= 500 ms, flip angle 55°; TE1=2.83 ms, TE2=5.29 ms) with similar slice and voxel parameters to the EPI sequence was utilized to obtain phase and magnitude maps utilized for unwarping of B$_0$ distortions of the functional data.

Data analysis

Behavioral responses occurring within 1.3 s after the target letter were accepted as correct. A nonparametric two-way Friedman ANOVA was used to examine the main effects of load and interference on performance. Because the subjects were advised to emphasize accuracy over the speed of performance these analyses concentrated on the hit rate (HR) measures. The medians and standard errors of medians of HRs, and reaction times (RT) for the correct detections, were estimated using bootstrapping with the data resampled 100,000 times.

Neuronal bases of auditory WM were studied using an fMRI-informed MEG/EEG approach, analogous to our previous studies (Ahveninen et al., 2013; Huang et al., 2014). External MEG noise was suppressed and subject movements, estimated continuously at 200-ms intervals, were compensated for using the signal-space separation method (Taulu et al., 2005) (Maxfilter, Elekta-Neuromag, Helsinki, Finland). The MEG/EEG data were then downsampled (300 samples/s, passband 0.5–100 Hz). Epochs coinciding with over 150 μV EOG, 100 μV EEG, 3000 fT/cm MEG gradiometer, or 4 pT MEG magnetometer changes were excluded from further analyses. To calculate fMRI-guided depth-weighted ℓ_2 minimum-norm estimates (MNE) (Hämäläinen et al., 1993; Lin et al., 2006), the information from structural segmentation of the individual MRIs and the MEG sensor and EEG electrode locations were used to compute the forward solutions for all putative sources in the cortex using a three-compartment boundary element model (Hämäläinen and Sarvas, 1989). The shapes of the surfaces separating the scalp, skull, and brain compartments were determined

from the anatomical MRI data using FreeSurfer 5.1 (http://surfer.nmr.mgh.harvard.edu/). For whole-brain inverse computations, cortical surfaces extracted with FreeSurfer were decimated to ~5,000 vertices per hemisphere. The individual forward solutions for current dipoles placed at these vertices comprised the columns of the gain matrix ($\mathbf{A}$). A noise covariance matrix ($\mathbf{C}$) was estimated from the raw MEG/EEG data during a 20–200 ms pre-stimulus baseline during non-active periods of the vigilance task. These two matrices, along with the source covariance matrix $\mathbf{R}$, were used to calculate the MNE inverse operator $\mathbf{W} = \mathbf{RA}^{\mathrm{T}} (\mathbf{ARA}^{\mathrm{T}} + \mathbf{C})^{-1}$.

To obtain an fMRI prior, *i.e.*, an fMRI-weighted source covariance matrix, each vertex point in the cortical surface was assigned an fMRI significance value using FreeSurfer-FSFAST 5.1. Individual functional volumes were motion corrected, unwarped, coregistered with each subject's structural MRI, intensity normalized, resampled into cortical surface space, smoothed using a 2-dimensional Gaussian kernel with an FWHM of 5 mm, and entered into a general-linear model (GLM) with the task conditions as explanatory variables. The fMRI weighting was set to 90%. That is, diagonal elements in $\mathbf{R}$ corresponding to vertices with below-threshold ($P < 0.05$, all conditions vs. baseline) significance values were multiplied by 0.1. The fMRI prior was defined for each subject as the union of the group average and their individual contrast.

One-second MEG/EEG raw data epochs, ranging from 400 ms before to 600 ms after the onset of the filler sounds, were multiplied by the inverse operator $\mathbf{W}$ and noise normalized to yield the estimated source activity as a function of time (Lin et al., 2006). The trial-specific source estimates were averaged within *a priori* regions-of-interest (ROI), including one AC and one DLPFC ROI per hemisphere that were determined based on the group results of our recent publication (Huang et al., 2013) (Figs. 2, 4). The AC ROIs included locations showing significant fMRI response suppression due to auditory WM load or interference in superior temporal locations, as determined based on Desikan et al. (2006). The DLPFC ROIs, which were used only for the connectivity analyses, were determined based on the significant (positive) main effect of WM load in (Huang et al., 2013). In ROI-to-ROI connectivity analyses, the waveform signs of sources were aligned based on the surface-normal orientations to avoid phase cancellations. Pre-stimulus oscillatory power was estimated in a 400-ms window preceding the filler sounds between and within the Q–A memory periods, during low and high-load conditions, using a fast Fourier transform (FFT), implemented in MNE-Python (Gramfort et al., 2013). In pre-stimulus power analyses, our hypotheses considered alpha-band (7–15 Hz) oscillations, whose role in WM and distracter suppression is supported by several studies in other sensory domains. Analyses at the beta (15–30 Hz) and gamma (30–100 Hz) bands are, however, also cursorily reported.

To compute dynamic functional connectivity estimates between the DLPFC and AC ROIs, 1.6 s trial epochs (600 pre-stimulus baseline) were convoluted with a Morlet Wavelet (width 5 cycles, 5–30 Hz, 1-Hz intervals). Phase locking between the four ROIs was then calculated at each time point using the pairwise phase consistency (PPC) measure (Vinck et al., 2010), which is not biased by the number of available trials.

Trial-based changes in AC-ROI activation patterns, pooled to bins of three raw-data time points, and their correlations with band-normalized pre-stimulus oscillatory power were analyzed using generalized linear mixed effects models (GLME, Matlab 9.0). These analyses modeled individual epochs in each task condition/subject, and brought longitudinal temporal dependencies into the statistical model, which allowed controlling for autocorrelations and other biases in aggregated estimates (Baayen et al., 2008; Baayen and Milin, 2010). The dependent variable, the amplitude of ROI-average responses, was modeled separately for each time bin, using a restricted maximum pseudo likelihood GLME with fixed effects for the pre-stimulus power, memory load (low vs. high), level of interference, and the trial number, and random effects for the intercept grouped by the subject identity. The GLME also controlled all possible interactions between the pre-stimulus power, memory load, and interference. We controlled for potential biases caused by non-normality and homogeneity by examining the model residuals against fitted values. Trial-to-trial autocorrelations were examined by determining the autocorrelation function for the residuals. Statistical significances were determined by comparing the sum t statistics of continuous positive or negative clusters of at least three consecutive time points (cluster-forming threshold $p<0.05$) derived from the initial GLME to a surrogate distribution calculated based on 5000 within-subject permutations. Comparisons of ROI-to-ROI time-frequency representations (TFR) of PPC were conducted using the Fieldtrip cluster-based randomization test (cluster-forming threshold $p<0.05$) (Maris and Oostenveld, 2007). As for the active Q–A periods, only periods preceding a correctly identified target were considered.

Results

The behavioral results validated our assumptions of the WM task manipulations: The hit rate decreased significantly as a function of increasing WM load (Friedman ANOVA, $\chi^2=16.41$, $p<0.001$) and attentional interference (Friedman ANOVA $\chi^2=13.99$, $p<0.001$). The median $\pm$ standard error of median hit rates was 94% $\pm$ 1% during the vigilance condition, 88% $\pm$ 2% under low WM load, 81% $\pm$ 4% under high WM load, 83% $\pm$ 5% during blocks with low WM load and interference, and 72% $\pm$ 6% during blocks with high WM load and interference. The median $\pm$ standard error of median reaction times was 478 $\pm$ 28 ms during the vigilance condition, 478 $\pm$ 37 ms under low WM load, 472 $\pm$ 46 ms under high WM load, 483 $\pm$ 12 ms under low WM load and interference, and 467 $\pm$ 24 ms under high WM load and interference.

Figs. 2a–c show the results of a GLME that analyzed the effects of WM load on auditory activations to filler sounds during the active Q–A memory periods and those during passive non-maintenance periods (Fig. 2, "other fillers"). The results suggest that the high vs. low memory load decreased negative MEG/EEG dipolar currents to filler sounds, which occurred during the active Q–A memory period, in the left AC. The effect was statistically significant 250–300 ms after stimulus onset (cluster-based GLME randomization test, $p<0.01$; Fig. 2c). Importantly, no similar effects emerged during the periods between the active Q–A periods. GLME model diagnostics during the latency showing the largest initial t-value is shown in Fig. 2c, which demonstrates that the residuals had no relationship with the fitted values, that the residuals were roughly normally distributed, and that there were no significant autocorrelations of the residuals. Individual-level estimates of the effects of WM

load on left AC activations to filler sounds are shown in Inline Supplementary Material (Suppl. Fig. 1).

To examine the hypothesis that dealing with irrelevant sounds reflects "active suppression" through "late selection" mechanism we conducted an interregional functional connectivity analysis between two frontal ROIs (the left and right DLPFC), which had been determined, analogously to the AC ROIs, based on a significant fMRI group effect by WM load in our previous publication (Huang et al., 2013). The main result (Fig. 2d) of this analysis suggests that WM load significantly increases 10–14 Hz alpha PPC between the left DLPFC and left AC at around 200–400 ms after the filler-sound onset ($p<0.01$, cluster-based randomization test). This effect, which was significant only during the active Q–A periods, was thus highly consistent with the timing of response suppression in the left AC (Fig. 2c). Individual-level data of left DLPFC–AC connectivity differences during high vs. low WM load are shown in Inline Supplementary Material (Suppl. Fig. 2). No significant WM-related PPC effects were observed between the right DLPFC and ACs and between the left DLPFC and right AC.

The load of *auditory* WM is hard to manipulate without changing the stimulus sequence, including the relative numbers of events across the conditions. To control for the resulting inherent differences in the low vs. high-load sound sequences, we conducted two alternative GLME analyses to verify the load-related suppression in left AC during the active Q–A periods (Fig. 3). The first control GLME considered only the first filler events after each Q. The second consisted of the main model with an additional factor controlling the type of the event that was immediately before each filler sound (see Fig. 1). The results of these control analyses were very highly similar to the main analysis in Fig. 2c, which supports the conclusion that the result reflects a true load effect instead of a sequential bias.

We then examined hypotheses regarding the role of AC alpha activations in suppression of irrelevant sounds during WM processing. The results of GLME models that tested the correlation between pre-stimulus oscillatory power and post-stimulus activation time courses are shown in Fig. 4. Similar to the load-related analyses (Figs. 2–3), significant correlations between response amplitudes and pre-stimulus alpha power were evidenced only during the active Q–A periods. High pre-stimulus alpha correlated with suppressed surface negative responses to filler sounds in the left AC 190–220 ms after the filler sound onset (cluster-based GLME randomization test, $p<0.05$). In the right AC, an analogous significant suppression of surface negativity was observed at 100–160 ms (cluster-based GLME randomization test, $p<0.05$). GLME model diagnostics during the latencies showing the largest cluster-forming t-values (Fig. 4c) demonstrate that the residuals had no relationship with the fitted values, that they were roughly normally distributed, and that there were no significant autocorrelations. Individual-level estimates of the effects of pre-stimulus alpha on AC activations to filler sounds are shown in Inline Supplementary Material (Suppl. Fig. 3).

Finally, we also analyzed correlations between pre-stimulus oscillations at the beta and gamma ranges and post-stimulus filler responses. No significant correlations were observed in these analyses.

Discussion

Our GLME analyses suggest that increased contextual maintenance load during an auditory WM task suppresses left AC activations at 250–300 ms after irrelevant phonetic "filler" stimuli. Interestingly, this effect coincided with increased alpha-range functional connectivity, quantified as 10–14 Hz PPC, between the left DLPFC and left AC ROIs. Although the load-related response suppression to irrelevant sounds, *per se*, would be consistent with a "filtering" hypothesis, the late time window of these suppression effects and the increasing DLPFC–AC functional connectivity are more consistent with a late selection or "active suppression" hypothesis. Finally, our results also demonstrate a trial-to-trial correlation between pre-stimulus alpha oscillations and post-stimulus suppression of AC responses to irrelevant filler sounds.

The way auditory WM demands affect auditory processing has remained unclear. Certain recent studies that employed intermodal distraction paradigms suggest that high visual WM load reduces auditory distractibility (Halin et al., 2015; SanMiguel et al., 2008). This could be speculated to mean that under high WM load, auditory distracters become filtered already at an early processing stage (*i.e.*, "early selection"). However, evidence from ERP (Berti and Schroger, 2003; Muller-Gass and Schroger, 2007) and behavioral (Dalton et al., 2009) studies that utilized purely auditory paradigms supports an alternative view that distracter suppression occurs at a later stage and requires active control (Berti and Schroger, 2003; Dalton et al., 2009). At the same time, recent behavioral studies also suggest that during auditory WM, irrelevant speech sounds are suppressed only after initial feature processing (Ellermeier et al., 2015; Wöstmann and Obleser, 2016). The present results seem to be more consistent with the latter group of ERP and behavioral studies: The aspect of left AC time course that was statistically significantly modulated by the contextual WM maintenance load was relatively late, clearly after the P50/N1/P2 pattern that is thought to reflect the emergence of conscious sound percept and the initiation of stimulus encoding in ACs (Parasuraman and Beatty, 1980). The latency of this effect would be more consistent with the "active suppression" hypothesis.

We also found evidence of load-related increases of alpha-band PPC between the left DLPFC and AC, which coincided with the suppression of AC responses to irrelevant sounds during the Q–A periods. Previous visual fMRI studies suggest that, in contrast to the more ventral aspects of PFC that deal with proactive interference (Jonides and Nee, 2006), DLPFC might support active WM maintenance by gating the incoming irrelevant stimuli during the retention delay period (Postle, 2005; Sakai et al., 2002). Notably, auditory ERP studies on neurological patients suggest that DLPFC gates already the earliest (10–50 ms) inputs to primary AC (Knight et al., 1989), providing a potential mechanism of "early selection". However, the present load-related suppression and connectivity effects occurred at a relative late stage of AC processing. This could mean that the DLPFC gating systems are able to regulate different stages of information processing depending on the task goals. That is, here, the subjects were given an instruction that required monitoring of the sound stream also during the Q–A periods. This involved, not only the easier task of counting the fillers, but also being prepared for occasional intervening Q or A events. (Note that the task instruction was similar before interference and non-interference task blocks.) Much like in

our introductory conversation example, it would not have been beneficial for the subject to completely halt auditory processing during the maintenance periods. We thus speculate that the increased alpha-range connectivity between the left DLPFC and AC is related to top-down modulation of post-perceptual processing, that is, after the subject becomes aware of the phonetic event onset (noting the 400-ms sound duration). The time window of this modulation is consistent with the period during which sensory representations have been suggested to be consolidated to more lasting WM traces (Chun and Potter, 1995; Jolicoeur and Dell'Acqua, 1998). A working hypothesis for future studies is that during auditory WM tasks, in which the competing events are embedded within the same sound stream, distracters are suppressed through a mechanism that allows initial evaluation but modulates consolidation of representations that would cause mnemonic interference at later WM processing stages.

The interpretation that the DLPFC–AC connectivity is related to AC suppression is consistent with previous theories that suggest a primarily inhibitory role for alpha-band connectivity patterns. For example, a recent auditory study suggested that alpha-range connectivity increases between frontal and ACs ipsilateral to the ear that the subject anticipates to be stimulated by a relevant event (Müller and Weisz, 2012). However, because the present oscillatory phase locking analyses do not provide information of the directionality of information flow, future studies using techniques such as transcranial magnetic stimulation (TMS) are needed to causally verify these interpretations.

It is important to note that in previous MEG and EEG studies of auditory selective attention, evidence for "early selection" has been obtained by using, for example, dichotic listening paradigms, in which the irrelevant stream contains little or no useful information (Woldorff et al., 1993). Because the present task contained relevant information also during the Q–A maintenance periods, our results do not fully exclude the possibility that "early selection" mechanisms could also be utilized during auditory WM. Another potential source of bias is that the number of filler sounds and their specific serial positions during the Q–A periods were slightly different between the high and low load conditions. This concern is mitigated by the robustness of GLME approaches in analyses with unequal numbers of observations (Baayen, 2008), as well at the fact that our control GLME models, which also considered the type of event that preceded each filler sound, made the overall results even clearer than the initial analysis. Another level of confirmation is offered by the aforementioned fact that the load-related suppression was also evident in analyses considering only the first fillers after each cue stimulus.

The present task consisted of 90 s blocks, during which the subject was prepared to be maintaining the task set for either short or longer series of events. At the event of each Q stimulus that occurred in the high-load blocks, the subject could be assumed to have "pre-allocated" more memory resources for the task set and sound maintenance than in the low-load blocks. Such an anticipatory strategy would make a lot of sense in this task, because the instruction was identical for high-load blocks with or without interference. A pre-allocation strategy could help explain why the difference between the high and low load conditions was also evident in analyses limited to the first filler sounds following the cues in low and high-load Q–A periods.

Our results also suggest that suppression of AC responses to irrelevant sounds during WM load is predicted by increased pre-stimulus alpha power. This negative correlation between pre-stimulus alpha power and sound-evoked responses is consistent with previous observations in the visual domain (Brandt and Jansen, 1991). Similar to the load-related effects, the present correlation effects were selective to task-irrelevant sounds that occurred during cue–target maintenance periods. One might thus speculate that AC alpha oscillations are related suppression of irrelevant inputs. This speculation receives support from a recent MEG finding that alpha power is increased in ACs ipsilateral to the ear receiving task-irrelevant stimulation during dichotic listening (Wöstmann et al., 2016).

A potential problem in the present alpha-correlation analysis, however, is the event-related AC response that follows sound onsets contains a lot of energy also at the alpha range and there is a risk that any fluctuations in "background" alpha that do not relate to either anticipation or active AC suppression produce false positive autocorrelations. Here, we attempted to control for these biases by using GLME approaches that also considered the temporal dependencies, including the exact timing of each trial and the type of preceding events, which we believe helped us regress out effects related to background oscillatory fluctuations, such as alpha-power fluctuations due to reduced alertness during prolonged experiments. With these controls in place, our GLME did, indeed, provide evidence of suppression of AC activities to irrelevant "filler" sounds during auditory WM processing. These effects were, however, earlier than the effects related to WM load modulations, specifically in the right AC (100–140 ms).

In conclusion, our GLME models suggest that activities elicited to irrelevant sound events are modulated at a relative late processing stage in the left AC under high contextual maintenance load of auditory WM. The load-related AC modulations coincided with increased phase synchronization of alpha oscillations between the left DLPFC and ACs. These effects could reflect a control mechanism that helps suppress short-term consolidation of irrelevant information, without affecting the perceptual encoding that is needed to determine whether a particular sound input is relevant or not. Our results also support the view that pre-stimulus alpha oscillations correlate with suppression of sound-evoked responses in ACs. Our results might help clinical investigation of disorders involving auditory WM dysfunctions, such as schizophrenia (Seidman et al., 2016).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Ahveninen J, Huang S, Belliveau JW, Chang WT, Hämäläinen M. Dynamic oscillatory processes governing cued orienting and allocation of auditory attention. J Cogn Neurosci. 2013; 25:1926–1943. [PubMed: 23915050]

Baayen, RH. Analyzing Linguistic Data: A Practical Introduction to Statistics Using R. Cambridge University Press; Cambridge: 2008.

Baayen RH, Davidson DJ, Bates DM. Mixed-effects modeling with crossed random effects for subjects and items. Journal of Memory and Language. 2008; 59:390–412.

Baayen RH, Milin P. Analyzing reaction times. Int J Psychol Res. 2010:12–28.

Berti S, Schroger E. Working memory controls involuntary attention switching: evidence from an auditory distraction paradigm. Eur J Neurosci. 2003; 17:1119–1122. [PubMed: 12653989]

Bonnefond M, Jensen O. Alpha oscillations serve to protect working memory maintenance against anticipated distracters. Curr Biol. 2012; 22:1969–1974. [PubMed: 23041197]

Brandt ME, Jansen BH. The relationship between prestimulus-alpha amplitude and visual evoked potential amplitude. Int J Neurosci. 1991; 61:261–268. [PubMed: 1824388]

Broadbent, DE. Perception and Communication. Pergamon Press; London: 1958.

Chun MM, Potter MC. A two-stage model for multiple target detection in rapid serial visual presentation. J Exp Psychol Hum Percept Perform. 1995; 21:109–127. [PubMed: 7707027]

Cooper NR, Croft RJ, Dominey SJ, Burgess AP, Gruzelier JH. Paradox lost? Exploring the role of alpha oscillations during externally vs. internally directed attention and the implications for idling and inhibition hypotheses. Int J Psychophysiol. 2003; 47:65–74. [PubMed: 12543447]

Dalton P, Santangelo V, Spence C. The role of working memory in auditory selective attention. The Quarterly Journal of Experimental Psychology. 2009; 62:2126–2132. [PubMed: 19557667]

Desikan R, Segonne F, Fischl B, Quinn B, Dickerson B, Blacker D, Buckner R, Dale A, Maguire R, Hyman B, Albert M, Killiany R. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage. 2006; 31:968–980. [PubMed: 16530430]

Deutsch JA, Deutsch D. Some theoretical considerations. Psychol Rev. 1963; 70:80–90. [PubMed: 14027390]

Ellermeier W, Kattner F, Ueda K, Doumoto K, Nakajima Y. Memory disruption by irrelevant noise-vocoded speech: Effects of native language and the number of frequency bands. J Acoust Soc Am. 2015; 138:1561–1569. [PubMed: 26428793]

Foxe JJ, Simpson GV, Ahlfors SP. Parieto-occipital approximately 10 Hz activity reflects anticipatory state of visual attention mechanisms. Neuroreport. 1998; 9:3929–3933. [PubMed: 9875731]

Gramfort A, Luessi M, Larson E, Engemann DA, Strohmeier D, Brodbeck C, Goj R, Jas M, Brooks T, Parkkonen L, Hämäläinen M. MEG and EEG data analysis with MNE-Python. Front Neurosci. 2013; 7:267. [PubMed: 24431986]

Halin N, Marsh JE, Sorqvist P. Central load reduces peripheral processing: Evidence from incidental memory of background speech. Scand J Psychol. 2015; 56:607–612. [PubMed: 26355647]

Hämäläinen M, Hari R, Ilmoniemi R, Knuutila J, Lounasmaa O. Magnetoencephalography-theory, instrumentation, and applications to noninvasive studies of the working human brain. Rev Mod Phys. 1993; 65:413–497.

Hämäläinen MS, Sarvas J. Realistic conductivity geometry model of the human head for interpretation of neuromagnetic data. IEEE Trans Biomed Eng. 1989; 36:165–171. [PubMed: 2917762]

Hillyard S, Hink R, Schwent V, Picton T. Electrical signs of selective attention in the human brain. Science. 1973; 182:177–180. [PubMed: 4730062]

Huang S, Chang WT, Belliveau JW, Hämäläinen M, Ahveninen J. Lateralized parietotemporal oscillatory phase synchronization during auditory selective attention. Neuroimage. 2014; 86:461–469. [PubMed: 24185023]

Huang S, Seidman LJ, Rossi S, Ahveninen J. Distinct cortical networks activated by auditory attention and working memory load. Neuroimage. 2013; 83:1098–1108. [PubMed: 23921102]

Jensen O, Kaiser J, Lachaux JP. Human gamma-frequency oscillations associated with attention and memory. Trends Neurosci. 2007; 30:317–324. [PubMed: 17499860]

Jensen O, Mazaheri A. Shaping functional architecture by oscillatory alpha activity: gating by inhibition. Front Hum Neurosci. 2010; 4:186. [PubMed: 21119777]

Jolicoeur P, Dell'Acqua R. The demonstration of short-term consolidation. Cogn Psychol. 1998; 36:138–202. [PubMed: 9721199]

Jonides J, Nee DE. Brain mechanisms of proactive interference in working memory. Neuroscience. 2006; 139:181–193. [PubMed: 16337090]

Klimesch W, Sauseng P, Hanslmayr S. EEG alpha oscillations: the inhibition-timing hypothesis. Brain Res Rev. 2007; 53:63–88. [PubMed: 16887192]

Knight RT, Scabini D, Woods DL. Prefrontal cortex gating of auditory transmission in humans. Brain Research. 1989; 504:338–342. [PubMed: 2598034]

Lavie N. Distracted and confused?: selective attention under load. Trends Cogn Sci. 2005; 9:75–82. [PubMed: 15668100]

Leiberg S, Lutzenberger W, Kaiser J. Effects of memory load on cortical oscillatory activity during auditory pattern working memory. Brain Res. 2006; 1120:131–140. [PubMed: 16989782]

Lin FH, Belliveau JW, Dale AM, Hämäläinen MS. Distributed current estimates using cortical orientation constraints. Hum Brain Mapp. 2006; 27:1–13. [PubMed: 16082624]

Maris E, Oostenveld R. Nonparametric statistical testing of EEG-and MEG-data. J Neurosci Methods. 2007; 164:177–190. [PubMed: 17517438]

Müller N, Weisz N. Lateralized auditory cortical alpha band activity and interregional connectivity pattern reflect anticipation of target sounds. Cereb Cortex. 2012; 22:1604–1613. [PubMed: 21893682]

Muller-Gass A, Schroger E. Perceptual and cognitive task difficulty has differential effects on auditory distraction. Brain Res. 2007; 1136:169–177. [PubMed: 17223092]

Näätänen, R. Attention and Brain Function. Lawrence Erlbaum; Hillsdale: 1992.

Näätänen R, Tervaniemi M, Sussman E, Paavilainen P, Winkler I. "Primitive intelligence" in the auditory cortex. Trends Neurosci. 2001; 24:283–288. [PubMed: 11311381]

Obleser J, Wostmann M, Hellbernd N, Wilsch A, Maess B. Adverse listening conditions and memory load drive a common alpha oscillatory network. J Neurosci. 2012; 32:12376–12383. [PubMed: 22956828]

Palva S, Palva JM. New vistas for alpha-frequency band oscillations. Trends Neurosci. 2007; 30:150–158. [PubMed: 17307258]

Parasuraman R, Beatty J. Brain events underlying detection and recognition of weak sensory signals. Science. 1980; 210:80–83. [PubMed: 7414324]

Pfurtscheller G. Induced oscillations in the alpha band: functional meaning. Epilepsia. 2003; 44(Suppl 12):2–8.

Postle BR. Delay-period activity in the prefrontal cortex: one function is sensory gating. J Cogn Neurosci. 2005; 17:1679–1690. [PubMed: 16269105]

Pulvermuller F, Kujala T, Shtyrov Y, Simola J, Tiitinen H, Alku P, Alho K, Martinkauppi S, Ilmoniemi RJ, Näätänen R. Memory traces for words as revealed by the mismatch negativity. Neuroimage. 2001; 14:607–616. [PubMed: 11506534]

Sakai K, Rowe JB, Passingham RE. Active maintenance in prefrontal area 46 creates distractor-resistant memory. Nat Neurosci. 2002; 5:479–484. [PubMed: 11953754]

SanMiguel I, Corral MJ, Escera C. When loading working memory reduces distraction: behavioral and electrophysiological evidence from an auditory-visual distraction paradigm. J Cogn Neurosci. 2008; 20:1131–1145. [PubMed: 18284343]

Seidman LJ, Breiter HC, Goodman JM, Goldstein JM, Woodruff PW, O'Craven K, Savoy R, Tsuang MT, Rosen BR. A functional magnetic resonance imaging study of auditory vigilance with low and high information processing demands. Neuropsychology. 1998; 12:505–518. [PubMed: 9805320]

Seidman LJ, Meyer EC, Giuliano AJ, Breiter HC, Goldstein JM, Kremen WS, Thermenos HW, Toomey R, Stone WS, Tsuang MT, Faraone SV. Auditory working memory impairments in

individuals at familial high risk for schizophrenia. Neuropsychology. 2012; 26:288–303. [PubMed: 22563872]

Seidman LJ, Shapiro DI, Stone WS, Woodberry KA, Ronzio A, Cornblatt BA, Addington J, Bearden CE, Cadenhead KS, Cannon TD, Mathalon DH, McGlashan TH, Perkins DO, Tsuang MT, Walker EF, Woods SW. Association of Neurocognition With Transition to Psychosis: Baseline Functioning in the Second Phase of the North American Prodrome Longitudinal Study. JAMA Psychiatry. 2016; 73:1239–1248. [PubMed: 27806157]

Taulu S, Simola J, Kajola M. Applications of the Signal Space Separation Method. IEEE Trans Signal Proc. 2005; 53:3359–3372.

Vinck M, van Wingerden M, Womelsdorf T, Fries P, Pennartz CM. The pairwise phase consistency: a bias-free measure of rhythmic neuronal synchronization. Neuroimage. 2010; 51:112–122. [PubMed: 20114076]

Wilsch A, Obleser J. What works in auditory working memory? A neural oscillations perspective. Brain Res. 2015

Woldorff MG, Gallen CC, Hampson SA, Hillyard SA, Pantev C, Sobel D, Bloom FE. Modulation of early sensory processing in human auditory cortex during auditory selective attention. Proceedings of the National Academy of Sciences of the United States of America. 1993; 90:8722–8726. [PubMed: 8378354]

Woldorff MG, Hackley SA, Hillyard SA. The effects of channel-selective attention on the mismatch negativity wave elicited by deviant tones. Psychophysiology. 1991; 28:30–42. [PubMed: 1886962]

Woldorff MG, Hillyard SA, Gallen CC, Hampson SR, Bloom FE. Magnetoencephalographic recordings demonstrate attentional modulation of mismatch-related neural activity in human auditory cortex. Psychophysiology. 1998; 35:283–292. [PubMed: 9564748]

Worden MS, Foxe JJ, Wang N, Simpson GV. Anticipatory biasing of visuospatial attention indexed by retinotopically specific alpha-band electroencephalography increases over occipital cortex. J Neurosci. 2000; 20:RC63. [PubMed: 10704517]

Wöstmann M, Herrmann B, Maess B, Obleser J. Spatiotemporal dynamics of auditory attention synchronize with speech. Proc Natl Acad Sci U S A. 2016; 113:3873–3878. [PubMed: 27001861]

Wöstmann M, Obleser J. Acoustic Detail But Not Predictability of Task-Irrelevant Speech Disrupts Working Memory. Front Hum Neurosci. 2016; 10:538. [PubMed: 27826235]

Highlights

- We studied distracter suppression under low/high auditory working memory (WM) load

- High WM load reduced auditory cortex (AC) activations 250–300 ms after distracters

- This involved increased prefrontal–AC alpha-band connectivity and AC alpha activity

- Under WM load, irrelevant sounds are suppressed through a "late", active process

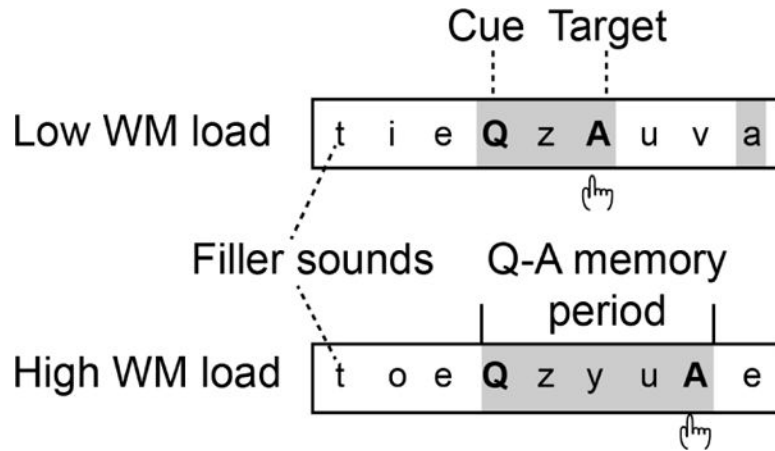- This allows initial screening but prevents consolidation of irrelevant information

Figure 1.
Stimuli and task. Subjects' task was to press a button whenever the target letter "A" occurred after the cue "Q" and a pre-specified number of interleaved filler sounds (average ISI=1 s, ±50 ms jitter). Other "non-target a" events were to be ignored. The fillers consisted of other letters of the alphabet. In "low WM load" blocks, "A" was a target when it was the second sound after a "Q". In "high WM load" blocks, "A" was a target when it was the fourth sound after a "Q". In certain "interference" task blocks, the Q–A periods could also include interfering cue and/or non-target "a" sounds (not shown). *We specifically examined whether and how the human brain suppresses responses to the task-irrelevant filler sounds that occur during vs. between the Q–A periods, during either low vs. high WM load task blocks.* The overall vigilance and compliance were controlled using a simple vigilance condition where "A" was the target when occurring immediately after a "Q" (not shown).
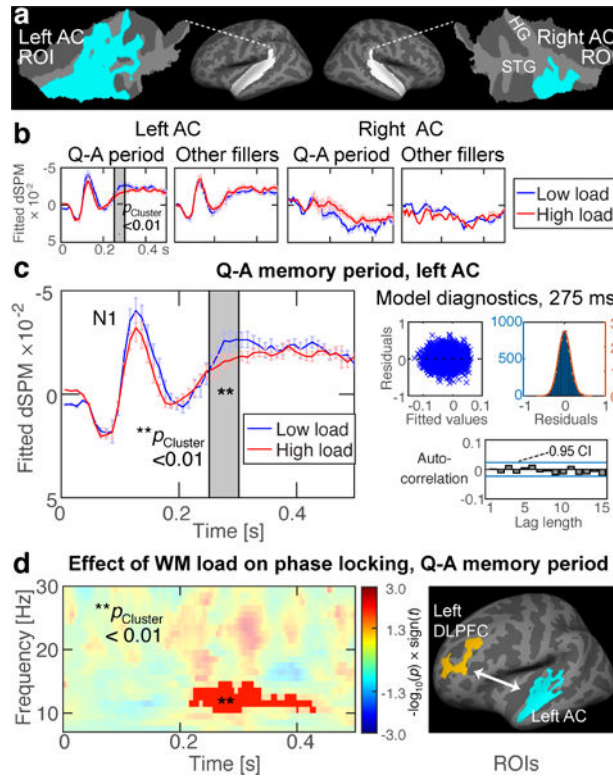
Figure 2.
Effect of WM load on AC responses to irrelevant "filler" sounds and frontal–AC connectivity. (**a**) AC regions-of-interest (ROI) (Huang et al., 2013) depicted atop flattened patches of superior temporal cortex. (**b, c**) GLME-modeled responses to filler sounds during low vs. high WM load, during active "Q–A periods" vs. passive periods (i.e., "other fillers"). In the left AC ROI, high WM load significantly suppressed negative activations to filler sounds. This occurred during the Q–A periods, 250–300 ms after sound onset (gray shading). No significant effects were observed in other comparisons. Error bars reflect the standard error of the mean, corrected for autocorrelations. The vertical axes ("fitted dSPM") reflect GLME-modeled values in arbitrary statistical units. (**c**) GLME diagnostics are shown at the within-cluster peak time point (275 ms). (**d**) High WM load enhanced alpha-band phase locking between the left DLPFC and AC ROIs (Huang et al., 2013) at around 200–400 ms after filler sounds ($p<0.01$, cluster-based randomization test). The timing is consistent with the suppression of sound-evoked responses in the left AC, shown in panel c.
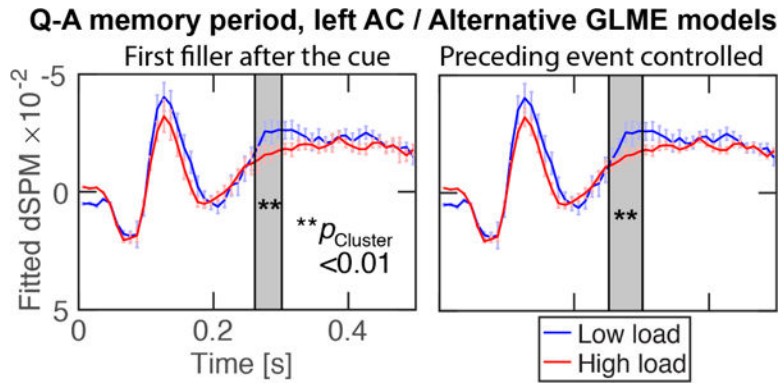
Figure 3.
Alternative GLMEs modeling the effect of WM load on left AC responses to irrelevant "filler" sounds. The first alternative GLME considered only the first filler response after each cue (left); the second was the main model with an additional factor for the event that preceded each filler. The results of these alternative analyses are highly consistent with the main model. The vertical axes ("fitted dSPM") reflect GLME-modeled values in arbitrary statistical units.
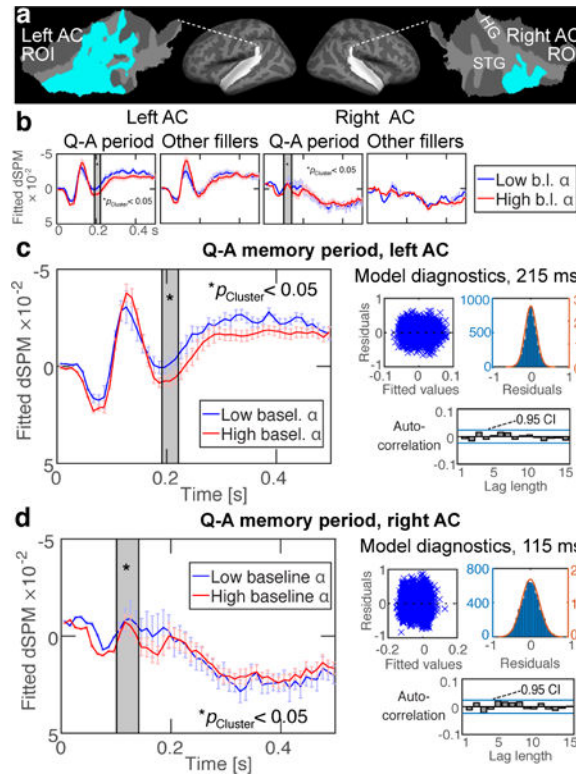
Figure 4.
Correlation between AC activations and pre-stimulus alpha power. (**a**) Regions-of-interest (ROI) depicted in flattened patches of superior temporal cortex. The ROIs were determined based on a previous study (Huang et al., 2013). (**b, c, d**) GLME results. High pre-stimulus alpha power correlated with suppression of post-stimulus negative activations to irrelevant "filler" sounds during the Q–A periods (**c**) in the left AC at 190–220 ms and (**d**) in the right AC at 100–140 ms (significant clusters shaded with gray). No significant effects were observed for filler sounds during the non-maintenance periods. The respective GLME diagnostics are shown for the within-cluster peak-effect time points. Error bars reflect the standard error of the mean, corrected for autocorrelations. The vertical axes ("fitted dSPM") reflect GLME-modeled values in arbitrary statistical units.