

# Selective Attention to Auditory Memory Neurally Enhances Perceptual Precision

Sung-Joo Lim,<sup>1,2</sup> Malte Wöstmann,<sup>1,2</sup> and  Jonas Obleser<sup>1,2</sup>

<sup>1</sup>Max Planck Institute for Human Cognitive and Brain Sciences, 04103 Leipzig, Germany and <sup>2</sup>Department of Psychology, University of Lübeck, 23562 Lübeck, Germany

Selective attention to a task-relevant stimulus facilitates encoding of that stimulus into a working memory representation. It is less clear whether selective attention also improves the precision of a stimulus already represented in memory. Here, we investigate the behavioral and neural dynamics of selective attention to representations in auditory working memory (i.e., auditory objects) using psychophysical modeling and model-based analysis of electroencephalographic signals. Human listeners performed a syllable pitch discrimination task where two syllables served as to-be-encoded auditory objects. Valid (vs neutral) retroactive cues were presented during retention to allow listeners to selectively attend to the to-be-probed auditory object in memory. Behaviorally, listeners represented auditory objects in memory more precisely (expressed by steeper slopes of a psychometric curve) and made faster perceptual decisions when valid compared to neutral retrocues were presented. Neurally, valid compared to neutral retrocues elicited a larger frontocentral sustained negativity in the evoked potential as well as enhanced parietal alpha/low-beta oscillatory power (9–18 Hz) during memory retention. Critically, individual magnitudes of alpha oscillatory power (7–11 Hz) modulation predicted the degree to which valid retrocues benefitted individuals' behavior. Our results indicate that selective attention to a specific object in auditory memory does benefit human performance not by simply reducing memory load, but by actively engaging complementary neural resources to sharpen the precision of the task-relevant object in memory.

**Key words:** alpha oscillations; attention to memory; electroencephalography; psychometric curve; retrocues

## Significance Statement

Can selective attention improve the representational precision with which objects are held in memory? And if so, what are the neural mechanisms that support such improvement? These issues have been rarely examined within the auditory modality, in which acoustic signals change and vanish on a milliseconds time scale. Introducing a new auditory memory paradigm and using model-based electroencephalography analyses in humans, we thus bridge this gap and reveal behavioral and neural signatures of increased, attention-mediated working memory precision. We further show that the extent of alpha power modulation predicts the degree to which individuals' memory performance benefits from selective attention.

## Introduction

Acoustic signals unfold in time as a series of fast-paced changes on a milliseconds time scale. Furthermore, acoustic signals of interest are most often intermixed with concurrent signals. To effectively perceive such transient and variable acoustic signals,

forming “auditory objects” (Griffiths and Warren, 2004) and maintaining them in memory is crucial. But can such internal representations of auditory objects be actively reselected from memory, and would such selection benefit auditory memory performance? An obvious candidate mechanism for such manipulations of memory content is selective attention, which enables effective encoding and maintenance of relevant information in working memory, at the expense of irrelevant information (Gazdaley and Nobre, 2012).

Previous visual working memory studies demonstrated that retroactive cues enable selective attention to task-relevant objects, thereby facilitating working memory performance (Sligte et al., 2008; Makovski et al., 2008; Pertzov et al., 2013). However, despite the acknowledged relevance of executive functions for auditory perception and its notorious challenges, our understanding is very limited with regard to the benefits that selective attention can provide when directed toward objects in auditory

Received July 13, 2015; revised Oct. 30, 2015; accepted Nov. 7, 2015.

Author contributions: S.-J.L. and J.O. designed research; S.-J.L. performed research; S.-J.L., M.W., and J.O. analyzed data; S.-J.L., M.W., and J.O. wrote the paper.

This work was supported by the Max Planck Society (Max Planck Research Group grant to J.O.). We are grateful to Dunja Kunke for her help in recording the data, and to Jöran Lepsien and Molly Henry for fruitful discussions at the design stage of this study.

The authors declare no competing financial interests.

Correspondence should be addressed to Dr. Sung-Joo Lim, Research Group “Auditory Cognition,” Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstrasse 1A, 04103 Leipzig, Germany. E-mail: sungjoo@cbs.mpg.de.

DOI:10.1523/JNEUROSCI.2674-15.2015

Copyright © 2015 the authors 0270-6474/15/3516094-11\$15.00/0

memory (Shinn-Cunningham, 2008). The neural mechanisms of retrospective attention implicated thus far are relatively specific to visual processing (but see Backer and Alain, 2012, 2014; Backer et al., 2015), and even those are still a matter of debate (see Souza et al., 2014).

The present study narrows the gap between evidence from visual retrospective attention studies and the inherently variable auditory objects by closely matching the requirements of previous visual retroactive-cue experiments in an auditory paradigm. Here, listeners encoded two easily categorizable speech syllables into memory (emulating the use of different visual objects) and were then cued to direct their attention to lower-level (pitch) information of one of these objects (emulating a visual object feature such as color or orientation). Furthermore, the current study adapted a psychophysical modeling approach established in the visual literature (Zhang and Luck, 2008; Bays and Husain, 2008; Murray et al., 2013) to obtain a fine-grained measure of memory performance in regard to the representational precision of objects in working memory.

Here, we focus on the underlying neural mechanisms of retrospective auditory attention. Attention-induced modulation of neural activity may reflect an enhancement of representational precision of the attended object in memory. In contrast, reduced modulation of neural activity, in line with visual-modality findings (Kuo et al., 2012), would suggest an attention-induced removal of unattended objects from memory. While both of these mechanisms can account for the facilitatory role of retrospective attention, it is unclear which of these postulated mechanisms is implemented by auditory attention. With the high temporal resolution of electroencephalography (EEG), we examine the effect of retrospective attention on the processing of a retroactive cue, the orientation of attention to one of two items in memory, and the ensuing retention of an item in memory.

Two candidate neural signatures of selective attention to auditory working memory in human EEG are conceivable: First, the magnitude of slow cortical potentials such as the contingent negative variation (CNV; Walter et al., 1964; Loveless and Sanford, 1975) reflects the amount of attention allocated in a task (Chennu et al., 2013; Wöstmann et al., 2015a). Moreover, the retention-related sustained anterior negativity may be a relevant component as its magnitude varies with auditory working memory load (Guimond et al., 2011; Lefebvre et al., 2013). Second, modulations of neural alpha ( $\sim 10$  Hz) oscillatory power are closely tied to selective attention and working memory load. Enhanced alpha power reflects greater demand on selective attention (Weisz et al., 2011; Wöstmann et al., 2015b) and/or higher memory load (Jensen et al., 2002; Tuladhar et al., 2007; Obleser et al., 2012), presumably through inhibition of task-irrelevant neural processes (Klimesch et al., 2007; Jensen and Mazaheri, 2010; Strauß et al., 2014a). However, evidence on how alpha power links to retrospective attention is sparse and restricted to the visual modality (Manza et al., 2014; Poch et al., 2014; Wallis et al., 2015). So far, a single auditory study reported alpha power modulations reflecting directing attention to memory, but their relation to the mechanisms underlying retrospective attentional benefits has remained unclear (Backer et al., 2015).

If retrospective attention facilitates memory performance by actively enhancing representational precision of the attended auditory objects, we would expect increased neural responses, such as the CNV and alpha power, reflecting increased attentional demands to retain precise memory representations. However, reduced neural responses would be expected if retrospective attention facilitates performance by removing unattended objects

from memory. Using psychophysical modeling and model-based analysis of EEG signals, we aim to reconcile the two potential mechanisms underlying auditory retrospective attention.

## Materials and Methods

### Participants

Thirty-nine (27 females, 12 males) native German speakers were recruited from the Max Planck Institute's participant database. Nineteen participants took part in the behavioral experiment, and  $n = 20$  others participated in the EEG experiment. All reported normal hearing and no histories of neurological disorders. Participants gave informed consent and received payment for the experimental time (7€ per hour). The study procedure was approved by the local ethics committee (University of Leipzig, Leipzig, Germany).

### Stimuli

Two syllable categories, /da/ and /ge/, were used in the experiment. Each syllable category consisted of six naturally varying tokens, spliced from three different utterances of two German words (/da/: "Dahlie," "Daten"; /ge/: "gegen," "gelen"). All utterances were recorded by a native German female speaker in a sound-attenuated booth and digitized at 44.1 kHz. Syllable tokens were truncated to be 200 ms in duration and edited with 3 ms linear onset and 30 ms offset ramps.

Four out of the six tokens for each category served as to-be-probed syllables presented during encoding. Given each syllable token, a set of eight probe stimuli was generated with parametrically varied pitch. To this end, the fundamental frequency ( $F_0$ ) was manipulated in eight steps:  $\pm 0.125$ ,  $\pm 0.375$ ,  $\pm 0.75$ , and  $\pm 1.25$  semitones relative to the to-be-probed syllable token. The average  $F_0$  of the /da/ stimuli was 162.2 Hz (range, 157.8–168.6) and that of the /ge/ stimuli was 176.5 Hz (range, 170.1–180.5).

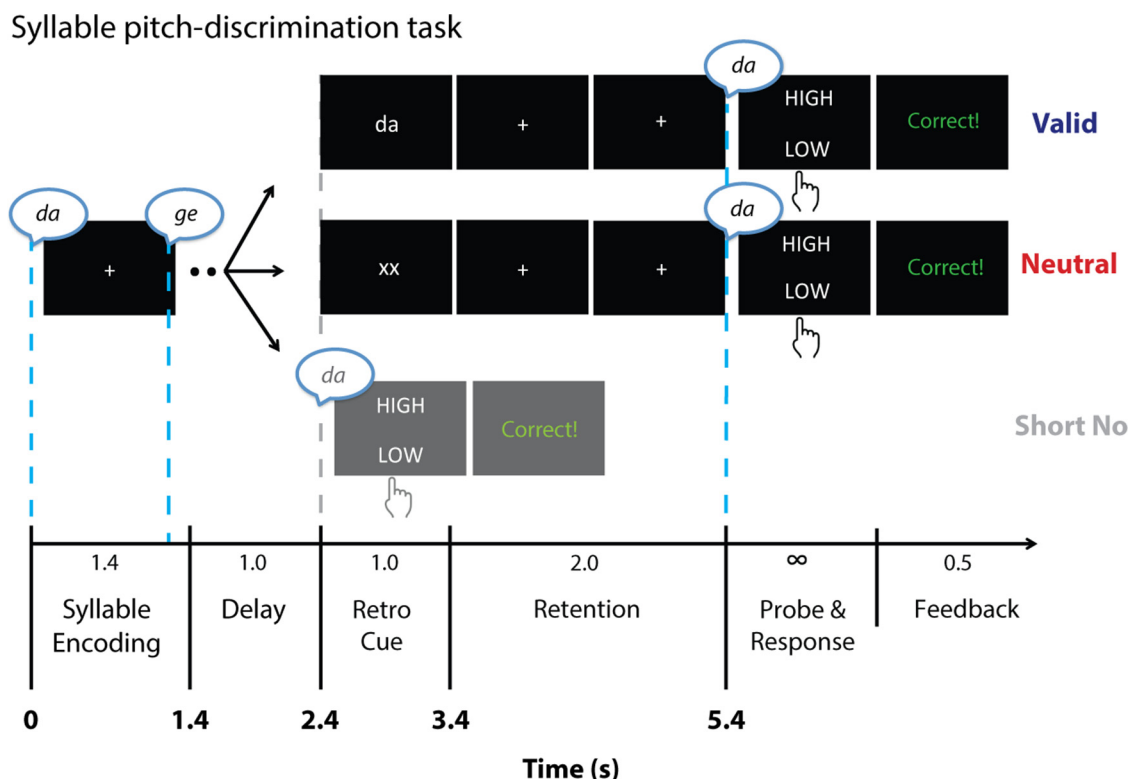
To increase acoustic variability beyond the fixed set of /da/ and /ge/ tokens, we created an additional 36 stimulus tokens for each syllable category. These stimuli were presented during encoding, but they were never probed for detecting pitch change. For this, the syllable tokens used to create probe stimuli (see above) were manipulated with pitch changes of  $\pm 0.5$  and  $\pm 0.625$  semitones. Also, the remaining two utterances of each syllable (/da/ and /ge/) recorded by the same speaker were used to serve as unprobed stimuli. The pitch of these utterances was manipulated with a maximum change of  $\pm 1.25$  semitones for a given syllable token. This manipulation range was restricted so that the  $F_0$ s of the unprobed syllables were variable, yet remained within the task-relevant  $F_0$  range of the set of probed syllable stimuli. This was to ensure that to-be-probed and unprobed syllables were not discriminable based on  $F_0$  during the encoding of the two syllables. On average, unprobed /da/ and /ge/ syllables had  $F_0$  values of 162.6 Hz and 175 Hz, respectively.

$F_0$  manipulation was accomplished with Praat version 5.3. All tokens were normalized to equivalent root-mean-squared amplitude (dB full scale).

### Task design and experimental procedure

Participants performed a syllable pitch discrimination task implemented within a retroactive cueing paradigm. The trial structure of the main task is illustrated in Figure 1. In each trial, participants heard /da/ and /ge/ syllables (0.2 s duration of each syllable) presented in a random order separated by 1 s silence interval in the encoding phase (i.e., 1.4 s total). This encoding phase was followed by a 1 s delay, during which participants maintained the two syllables in working memory. After this delay (i.e., 1 s after the offset of the second syllable), a visual retrocue was displayed on the screen for 1 s. After an additional 2 s delay following the cue, an auditory probe stimulus was presented. At the end of each trial, participants judged whether the pitch of the probe syllable was higher or lower compared to the same syllable category sound presented during the encoding phase (i.e., beginning of the trial). For instance, if participants heard /da/ stimulus as a probe, then the pitch of this probe should be compared to the /da/ sound presented in the encoding phase. After providing a response to a given probe, participants received visual feedback for 0.5 s.

There were three types (conditions) of trials in the experiment. One is called the "valid" retrocue trial, in which a visual retrocue was presented



**Figure 1.** Trial structure of the syllable pitch discrimination task. On each trial across all three retrocue trial conditions, participants heard a sequence of two syllables (0.2 s for each syllable) separated by a 1 s silence interval during the encoding phase (i.e., 1.4 s in total). The trial condition manipulation occurred after a 1 s delay following the encoding phase. In the valid and neutral conditions, participants were presented with either an informative “valid” or noninformative “neutral” visual retrocue (note that no “invalid” or otherwise misleading cues occurred). After the following 2 s stimulus-free retention phase, participants heard a probe stimulus and judged whether the pitch of the probe syllable was higher or lower compared to the same category syllable heard during encoding (for details, see Materials and Methods). In the short no-cue trials, participants heard a probe stimulus instead of seeing a visual retrocue as in other cueing conditions.

to provide information about which of the two syllables would be probed. In these trials, participants were presented with either a written “da” or a written “ge” as a visual retrocue (verdana font; approximate visual angle, 1.72°) and heard a probe from the corresponding syllable category (note that no “invalid” or otherwise misleading cues occurred). Another type of trial, called the “neutral” retrocue trial, did not provide any useful retrocue information; participants only saw “xx” on the screen, indicating that either of the two encoded syllables could be probed. Thus, participants had to retain information about both syllables in working memory until hearing a probe.

A third trial type, called the “short no-cue” trials, only served as control for a potential detrimental effect of temporal delay from retention phase on recall performance. In these trials, an auditory probe syllable was presented at the time at which a visual retrocue was expected (2.4 s; Fig. 1; Makovski and Jiang, 2007; Makovski et al., 2008; Murray et al., 2013). Thus, whereas the valid and neutral retrocue trials assessed pitch change detection performance with 4 s delay after syllable encoding, the short no-cue trials assessed performance with a relatively short delay period (i.e., 1 s). Participants were unaware of the trial type and the to-be-probed syllable category until seeing a retrocue (in valid and neutral cue conditions) or hearing a probe (in the short no-cue condition).

A central fixation cross was present throughout the trial period except for during the visual retrocue, response prompt, and feedback screen. Participants went through a total of 16 blocks, and each consisted of 24 trials (i.e., eight probe steps by three retrocue conditions). Within each block, two syllable positions during encoding were equally probed. This way, we prohibited listeners to build any expectation about the probed syllable during the encoding phase. Thus, both syllables were equally important for the task across all trial types.

Before the main experiment, participants were briefly instructed about the experimental task. Participants first went through a practice session (18 trials total, 6 in each condition) only with probes with easily detectable pitch changes ( $\pm 1.25$  semitones). This session was given to ensure that participants understood the task. If the practice performance did not

reach >80% accuracy, the practice session was repeated. (This was the case for 11 of 39 participants.)

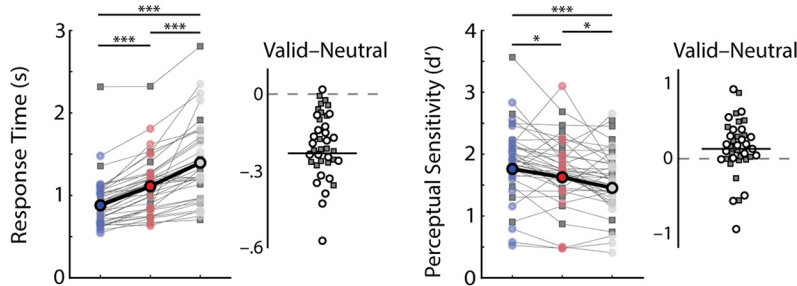
Experimental trials were controlled with Presentation software (Neurobehavioral Systems). Auditory stimuli were delivered via headphones (Sennheiser HD 25-SP II) at 50 dB above the individual’s sensation level (50 dB SL; sensation level was predetermined individually using this experiment’s stimuli and the method of limits). The behavioral experiment was conducted in a sound-attenuated booth. The EEG experiment was conducted in an electrically shielded sound-attenuated EEG booth. The same experimental design was used for both studies, except for a short self-paced break, which was inserted in between trials of the EEG session. For the behavioral experiment, trials were separated by 2 s inter-trial intervals. For the EEG experiment, trials started after 1 s delay interval following a self-paced break in between trials.

#### Behavioral data analysis

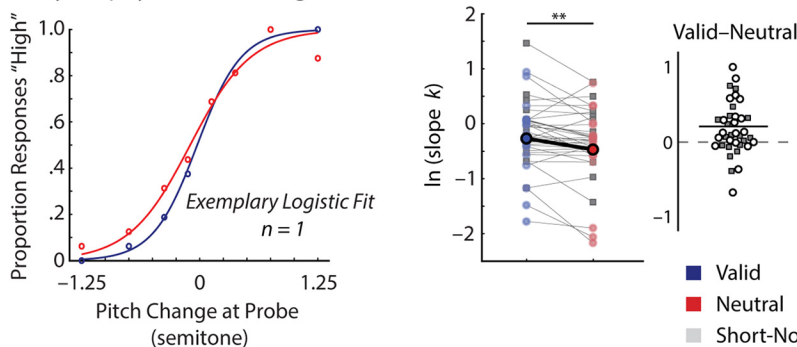
Since the experimental task design was identical for both behavioral and EEG studies, we analyzed the behavioral data across all participants ( $N = 39$ ).

**Behavioral measures.** Response times (RTs) relative to the onset of the probe syllable and performance accuracy were measured. All trials (correct and incorrect) were included in the analyses. To obtain a bias-free performance measure of perceptual sensitivity, each participant’s sensitivity to pitch change was calculated according to signal detection theory (Macmillan and Creelman, 2004). Our main interest was to contrast behavioral measures in the valid versus neutral retrocue condition. Moreover, the control (short no-cue) condition was included in the analysis of behavioral data only. The rationale for this was to examine whether performance is affected by duration of the retention period in the two retrocue conditions. For statistical analyses of RT and  $d'$  measures, we conducted two separate mixed ANOVAs with retrocue condition (valid, neutral, and short no cue) as a within-subjects factor

**A Behavioral Results**



**B Psychophysical Modeling Results**



**Figure 2.** Task performance and modeling parameter estimates as a function of retrocue condition. **A**, Behavioral performance of each retrocue condition including the short no-cue control condition is shown in RTs (left) and perceptual sensitivity ( $d'$ ; right). Data points connected by thin lines indicate individuals' performances across retrocue conditions. Gray squares and circles respectively indicate performances in the behavioral-only experiment participants ( $n = 19$ ) and the EEG participants ( $n = 20$ ). Larger dots connected with a bold line indicate mean performances of the EEG participants. Density plots illustrate the individuals' differences (valid vs neutral) in the respective measures. The behavioral-only and EEG participants are respectively indicated with gray squares and circles. The solid line indicates the mean difference (valid vs neutral) of the EEG participants. **B**, Parameter estimates of psychophysical modeling results. Left, An example of a single participant's logistic model fits is illustrated for the valid and neutral retrocue conditions. Each open dot represents the proportion of "high" responses as a function of pitch change relative to the reference syllable presented during the encoding phase. The lines represent model fits. Right, Individuals' and group average perceptual precision estimates (log-transformed slope,  $k$ ) of retrocue conditions and individuals' differences (valid vs neutral) of the precision estimates are shown with the same illustration scheme in **A**. \* $p < 0.05$ ; \*\* $p < 0.005$ ; \*\*\* $p < 0.0005$ .

and the experimental setting (behavioral-only vs EEG) as a between-subjects factor in SPSS (version 21).

**Psychophysical modeling.** Beyond the behavioral performance measures of RTs and  $d'$ , we estimated a more fine-grained, perceptual precision measure with a psychophysical modeling approach (Zhang and Luck, 2008; Bays and Husain, 2008; Murray et al., 2013). To quantify individual's perceptual precision in detecting the syllable pitch change, we fitted each participant's response patterns to the varying levels of F0 change that occurred at the probe. To this end, we used a nonlinear least squares curve-fitting procedure (lsqcurvefit function from MATLAB) with a logistic (sigmoid) function,  $y = 1/(1 + e^{-k(x - m)})$ , where  $y$  indicates the proportion of "high" responses,  $x$  indicates F0 change (in eight steps) at the probe relative to the encoded syllable in working memory,  $k$  indicates the slope, and  $m$  indicates the inflection point of the logistic function on the  $x$ -axis. The inflection point ( $m$ ) provides an estimate of response bias. The slope ( $k$ ) estimates the perceptual precision in the pitch change detection: the steeper the slope, the greater the perceptual precision (Fig. 2B; left).

With our main interest in the contrast between the two retrocue conditions (valid and neutral), psychophysical modeling estimates of slope ( $k$ ) and bias ( $m$ ) were analyzed using a mixed ANOVA with retrocue condition (valid vs neutral) as a within-subjects factor and the experimental setting (behavioral-only vs EEG) as a between-subjects factor. We report  $p$  values based on Greenhouse–Geisser corrected degrees of freedom in cases where the sphericity assumption was violated (Mauchly's test,  $p < 0.05$ ).

Any significant retrocue condition effects found in ANOVAs were followed up by *post hoc* paired samples  $t$  tests for each pair of retrocue conditions.

**EEG data acquisition and preprocessing**  
EEG data were continuously acquired from 66 electrodes (Ag–AgCl), including 61 scalp electrodes (Waveguard, ANT Neuro), one nose, and two mastoids (A1 and A2). The electrooculogram was acquired to record eye movements, with two electrodes placed horizontally to each eye and vertically to the right eye. A ground electrode was placed at the sternum. All impedances were set below 5 k $\Omega$ . The left mastoid (A1) served as reference during recording. The data were acquired with a sampling rate of 500 Hz and a hardware-implemented pass-band of DC to 135 Hz (TMS International).

Before EEG recording, we recorded individual electrode locations with the Polhemus FASTRAK electromagnetic motion tracker for source localization of EEG responses.

The data were preprocessed and analyzed with MATLAB using the FieldTrip toolbox (Oostenveld et al., 2011) and customized scripts. To observe responses of all events in the trial, the continuous data were divided into epochs of  $-2$  to  $6$  s relative to the trial onset (i.e., onset of a first syllable during encoding). Observing the whole trial epoch ensures inspecting for any spurious effects in the time period during which no condition effects are to be plausibly expected (i.e., before the onset of the retrocue). An independent component analysis was performed, and components relating to eye movements, electrical heartbeat, and noisy components were removed from the data (Debener et al., 2010). On average,  $14.25 \pm 3.43$  (mean  $\pm$  SD) of 61 components were removed. Moreover, epochs were removed if any scalp electrode showed an activity range of  $>200 \mu V$  within the  $-1$  to  $6$  s time window relative to the trial onset. Through this procedure,  $\sim 7\%$  of epochs were rejected on average per participant; this resulted, on average, in  $119.15 \pm 5.98$  (mean  $\pm$  SD) valid and  $120.4 \pm 6.37$  neutral trials for further statistical analyses.

**Event-related potentials**

The single-trial epoched EEG data were baseline corrected by subtraction of the mean amplitude in the time interval  $-0.3$  to  $-0.1$  s relative to trial onset. Single-trial data from  $0$  to  $5.4$  s (i.e., time window from trial onset to retention offset) were used to contrast evoked responses [event-related potentials (ERPs)] of the valid versus neutral retrocue trials with a multilevel statistical analysis (see Statistical analyses section, below).

**Time–frequency representations**

Time–frequency representations (TFRs) of each trial were computed by convolving the single-trial time-domain EEG data with a family of seven-cycle Morlet wavelets for frequencies of  $1$ – $40$  Hz (with  $1$  Hz resolution). This procedure was applied in  $10$  ms steps from  $-2$  to  $6$  s relative to trial onset. To avoid artifacts occurring at the edges (trial beginning and end) from time–frequency decomposition, we used a "reflection" approach (Cohen, 2014; van den Brink et al., 2014). This approach creates an extra buffer zone of no interest, containing only redundant time–frequency content, at the beginning and end of each trial by concatenating the time- and polarity-inverted (i.e., mirrored) EEG signal of the whole trial window. The increased epoch length accommodated the width of wavelet cycles especially in low frequencies. After the time–frequency decomposition, the mirrored EEG signal in the buffer zone was discarded, and only the original trial epoch was preserved. Baseline correction was applied to single-trial power estimates as a ratio of change relative to the average power estimate during the  $0.5$  s time window before trial onset.

### Statistical analyses of event-related potentials and time–frequency representations

Multilevel statistical analyses were performed for the ERPs and TFRs (Obleser et al., 2012). First, the single-subject-level statistical analyses of ERPs and TFRs of all trials (correct and incorrect) were performed on single-trial data from 0 to 5.4 s (i.e., time window from trial onset to retention offset). Contrast coefficients of the valid and neutral retrocue trials were respectively set to 0.5 and  $-0.5$  for independent samples regression coefficient  $t$  tests implemented in FieldTrip. For this analysis, the `ft_timelockstatistics` and `ft_freqstatistics` functions in FieldTrip were used for the ERP and TFR data, respectively. This analysis resulted in  $\beta$  weights of the retrocue condition contrast for each time–electrode data point for the ERPs and for each time–electrode–frequency data point for the TFRs.

Next, the group-level analysis was performed with a dependent samples  $t$  test to contrast the  $\beta$  weights from the subject-level analysis against zero. For the ERPs, the resulting  $\beta$  weights from the subject-level analysis were entered into the group analysis. For the TFRs, the group-level analysis was performed on the  $\beta$  weights of the frequency range from 1 to 40 Hz. A permutation test (1000 Monte Carlo random iterations) was performed with cluster-based control at a type I error level of  $\alpha = 0.05$  as implemented in FieldTrip. This analysis resulted in time–electrode and time–electrode–frequency clusters exhibiting significant retrocue condition differences in the ERPs and the TFRs, respectively. Note that for the ERP data with fine temporal resolution (500 Hz), time–electrode clusters in close proximity ( $\sim 120$  ms) exhibiting same direction of effect were collapsed; that is, the union of these clusters across time and electrodes was averaged.

### Brain–behavior correlations (model-based EEG analysis)

We further conducted correlational analyses to relate modulations in both neural measures (i.e., ERPs and oscillatory power) across the retrocue conditions to interindividual differences in the behavioral benefit from valid versus neutral retrocues. For the ERP data, we extracted average amplitude differences in evoked responses (valid vs neutral) from each of the clusters exhibiting a significant effect of retrocue condition. To quantify overall strength of ERP across broad regions of the scalp regardless of the polarity differences of potentials, evoked activity at each time point was expressed as global field power (GFP), the spatial SD across electrodes (Lehmann and Skrandies, 1980; O’Sullivan et al., 2015). We calculated the Spearman correlation of the retrocue condition difference in evoked activity ( $GFP_{\text{Valid}}$  vs  $GFP_{\text{Neutral}}$ ), and the differential performance on perceptual sensitivity ( $d'_{\text{Valid}}$  vs  $d'_{\text{Neutral}}$ ) and on perceptual precision ( $\ln k_{\text{Valid}}$  vs  $\ln k_{\text{Neutral}}$ ), separately.

For the TFR data, we focused on a model-based EEG analysis: Individual parameter estimates from psychophysical modeling were used as a regressor in a permutation-based statistical test across all time–frequency–electrode bins, to examine the relationship between the extent of oscillatory power modulations and of behavioral modulations by retrocue condition. For each participant, the retrocue-relevant modulation of perceptual precision was measured as difference between the log-transformed slope estimates of the valid and neutral retrocue trials ( $\ln k_{\text{Valid}}$  vs  $\ln k_{\text{Neutral}}$ ). This difference was regressed against the degree of individuals’ oscillatory power modulations within the frequency range of 1–40 Hz, that is, differences between oscillatory power of the valid and neutral retrocue averaged across trials ( $\text{power}_{\text{Valid}}$  vs  $\text{power}_{\text{Neutral}}$ ). A cluster-based permutation approach was used to find clusters of time points, electrodes, and frequencies showing significant correlations between modulations of oscillatory power and perceptual precision. Using a similar permutation-based approach, we also correlated the difference between overall behavioral perceptual sensitivities ( $d'_{\text{Valid}}$  vs  $d'_{\text{Neutral}}$ ) with the oscillatory power difference between conditions ( $\text{power}_{\text{Valid}}$  vs  $\text{power}_{\text{Neutral}}$ ).

### Source localization of time–frequency effects

To localize neural oscillatory effects found in the sensor-level analyses, we further conducted a source analysis. To this end, individual EEG electrode positions of each participant were coregistered with the standard MRI template surface (using affine transformation). The head

model was based on the FieldTrip’s boundary element method (Oostenveld et al., 2003). All data were re-referenced to the average reference, and individuals’ lead field matrices were then calculated with 1 cm grid resolution.

Source localization of oscillatory power modulations found in the sensor-level clusters was performed using dynamical imaging of coherent sources (Gross et al., 2001) and following the FieldTrip-implemented beamforming technique (Haegens et al., 2010; Obleser and Weisz, 2012; Obleser et al., 2012; Strauß et al., 2014b). In short, a spatially adaptive filter was derived from the cross-spectral densities (CSDs) for all sensors. The CSD matrix was computed using a multitaper fast Fourier transformation (FFT) on single trials. Based on the sensor-level alpha/beta power effects found during the cue and retention phases, frequency estimates were centered at 13 Hz ( $\pm 4$  Hz smoothing), and time windows of interest were respectively set to 2.6–3.6 s and 3.9–4.9 s. With the individual’s lead field and the CSDs of all data (across conditions and baseline), a common filter was constructed to source project alpha/beta power modulations of each trial in these two time windows. The spatial distribution of power of single trials was then computed as a relative power change against the averaged source-projected alpha/beta power during baseline,  $-1.0$  to  $0$  s relative to trial onset.

To localize the sensor-level condition effects for each time window of interest, we performed multilevel statistical  $t$  tests separately for the cue presentation and retention phases to contrast valid versus neutral conditions. The resulting  $t$  values were interpolated to the standard MNI space and projected onto a standard MNI brain (SPM 8). Note that this statistical testing was performed only to visualize source reconstruction of the condition effect found in the sensor-level analysis. Thus, no stringent cluster-level thresholding was applied.

We also aimed to source localize the alpha power modulation ( $\alpha_{\text{Valid}}$  vs  $\alpha_{\text{Neutral}}$ ) that would best predict individual’s perceptual precision modulation ( $\ln k_{\text{Valid}}$  vs  $\ln k_{\text{Neutral}}$ ). Based on the sensor-level cluster result, a multitaper FFT was centered at 9 Hz ( $\pm 2$  Hz smoothing) and spanned a time window of 3.2–4.2 s. A corresponding common filter was constructed, and the relative source-projected alpha power change against baseline was computed as above. Given the average source-projected alpha power difference (valid vs neutral), we performed a permutation-based analysis to localize the correlations between modulations of alpha power and perceptual precision. As above, the resulting  $t$  values were interpolated to and projected onto a standard MNI brain for visualization purposes.

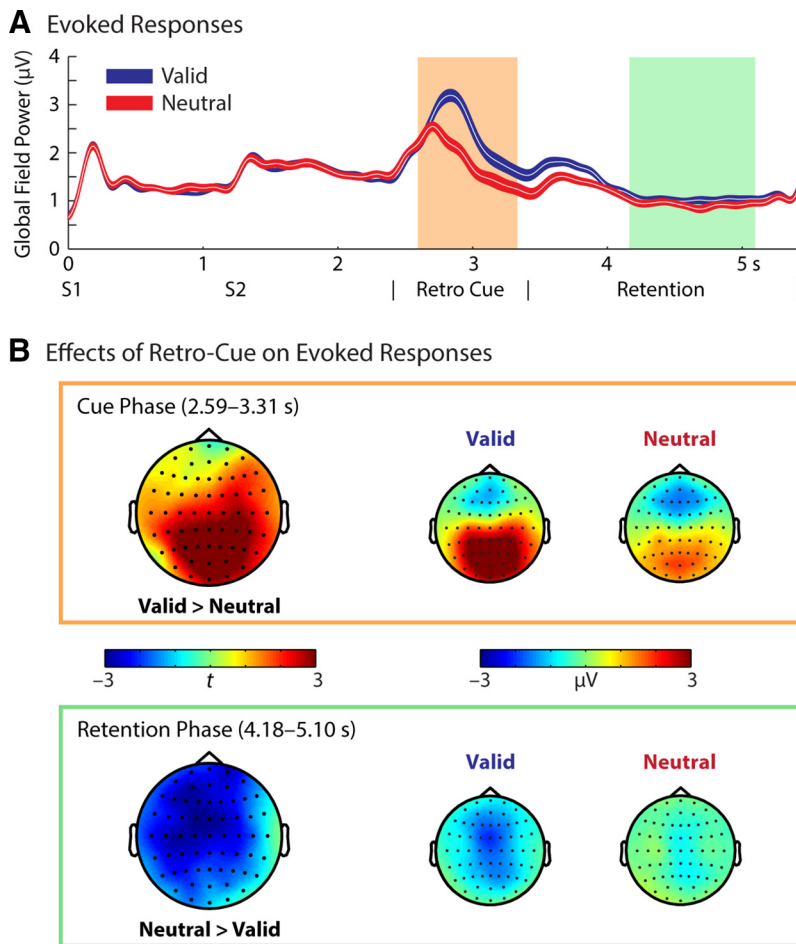
## Results

### Valid retrocues facilitate task performance

First, we analyzed whether participants’ RTs differed across different retrocuing conditions. A two-way mixed ANOVA revealed a significant main effect of retrocue condition ( $F_{(2,74)} = 81.79$ ; Greenhouse–Geisser  $\epsilon = 0.65$ ;  $p < 0.0005$ ;  $\eta_p^2 = 0.69$ ), but no main or interaction effects of experimental setting (both  $p$  values  $> 0.36$ ). As illustrated in Figure 2A (left), participants were significantly faster in judging the pitch of probe syllables with a valid retrocue compared to neutral retrocue ( $t_{(38)} = 9.87$ ;  $p < 0.0005$ ) or short no-cue trials ( $t_{(38)} = 10.26$ ;  $p < 0.0005$ ). Short no-cue trials yielded even longer response times than neutral cue trials ( $t_{(38)} = 6.87$ ;  $p < 0.0005$ ).

Similarly, a two-way mixed ANOVA on participants’ perceptual sensitivity (measured as  $d'$ ) indicated a significant main effect of retrocue condition ( $F_{(2,74)} = 11.61$ ;  $p < 0.0005$ ;  $\eta_p^2 = 0.24$ ), but no significant main or interaction effects related to the experimental settings (both  $p$  values  $> 0.19$ ). In Figure 2A (right), pairwise  $t$  tests revealed that participants’ pitch judgments were more accurate in trials with a valid retrocue than in neutral cue trials ( $t_{(38)} = 2.56$ ;  $p = 0.015$ ) or short no-cue trials ( $t_{(38)} = 5.20$ ;  $p < 0.0005$ ). Again, perceptual sensitivity in short no-cue trials was lower than in neutral cue trials ( $t_{(38)} = 2.18$ ;  $p = 0.036$ ).

Note that since the short no-cue trials yielded quantitatively and qualitatively different performance (slower responses, more



**Figure 3.** Effects of retrocues on evoked responses. **A**, Evoked responses shown as grand average of GFP, averaged across all (61) electrodes and EEG participants ( $n = 20$ ; low-pass filtered for illustration purposes only at 5 Hz with a sixth-order zero-phase shift Butterworth filter). The GFP collapsing across polarities in broad scalp regions (see topographical maps in **B**) is plotted to illustrate the overall neural activity. Shaded error lines indicate  $\pm 1$  within-subjects SE. Highlighted regions indicate the time period with a significant retrocue condition effect (valid vs neutral) from the multilevel statistical test. S1 and S2 denote two syllables presented during encoding. **B**, Topographical maps of the significant clusters for the valid versus neutral retrocue condition contrast (left) and evoked response amplitudes of each condition (right) in the time periods of the corresponding clusters. Highlighted electrodes belong to the corresponding clusters. Top, Topographical maps during the retrocue presentation phase. Bottom, Topographical maps during the retention phase.

errors) and had different trial timing compared to the trials with a retrocue (Fig. 1), we focused all ensuing psychophysical modeling and EEG analyses on comparing the valid against the neutral retrocue trials only.

Figure 2B illustrates the results of psychophysical modeling of perceptual precision and response bias, quantified by the  $k$  and  $m$  parameters of the logistic function, respectively (for details, see Materials and Methods). A two-way mixed ANOVA on the log-transformed slope ( $k$ ) of the logistic function fit revealed a significant main effect of cue condition ( $F_{(1,37)} = 8.90$ ;  $p = 0.005$ ;  $\eta_p^2 = 0.19$ ), but no significant main or interaction effects related to experimental settings (both  $p$  values  $> 0.3$ ). As predicted, perceptual precision in the valid retrocue condition was significantly higher compared to the neutral retrocue condition ( $t_{(38)} = 3.02$ ;  $p = 0.004$ ; Fig. 2B, right). For the bias parameter estimate ( $m$ ), a mixed ANOVA revealed no significant effect of cue condition ( $F_{(1,37)} = 0.03$ ;  $p = 0.87$ ;  $\eta_p^2 = 0.001$ ). One-sample  $t$  tests of the bias estimates against 0 (i.e., no bias) revealed that neither the valid nor the neutral retrocue conditions induced a significant bias to judge the probe pitch (valid,  $t_{(38)} = 1.87$ ,  $p = 0.07$ ; neutral,  $t_{(38)} = 1.60$ ,  $p = 0.12$ ).

### Valid and neutral retrocues differentially affect evoked responses

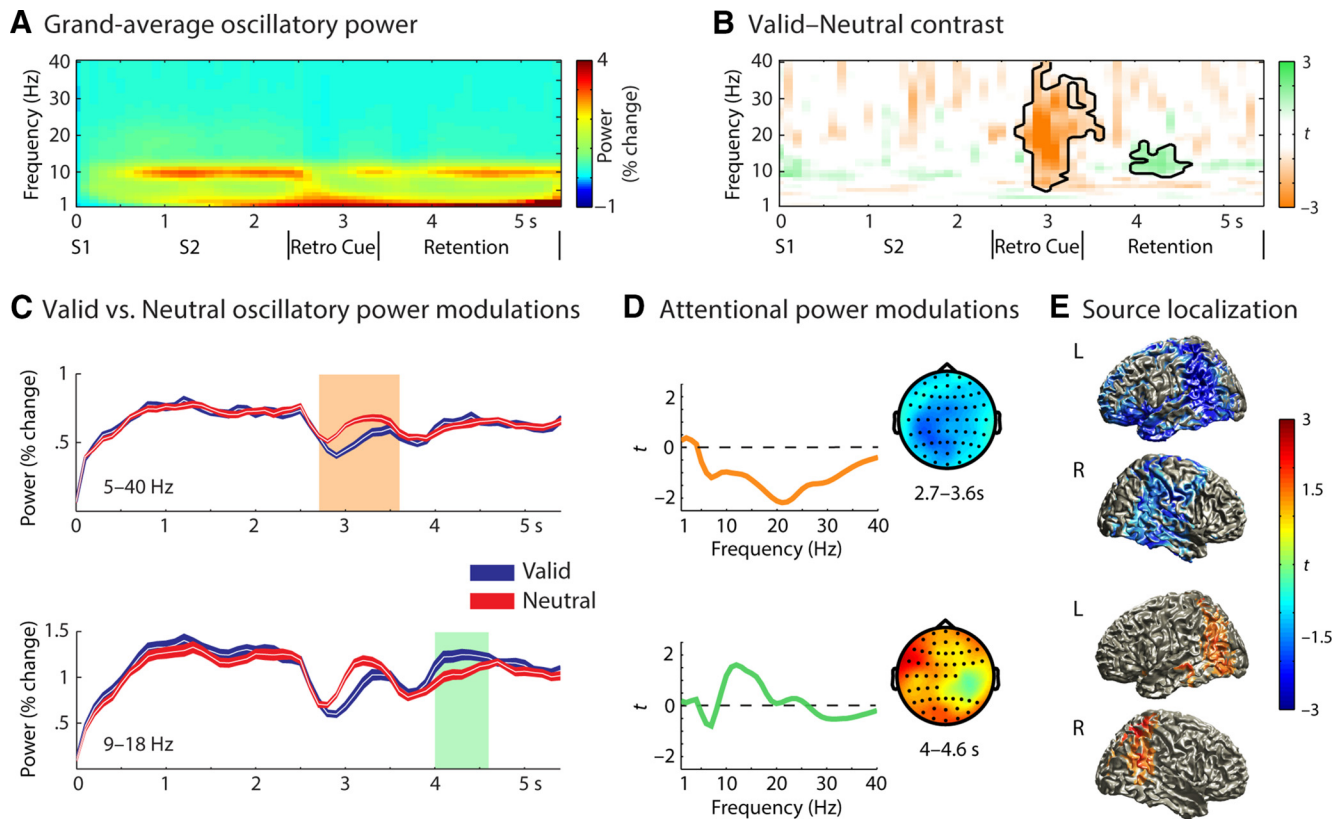
Figure 3A illustrates evoked responses throughout the trial period for the valid and neutral retrocue conditions, expressed as grand average of GFP, a measure of overall response strength across all scalp electrodes. Expectedly, since participants were not aware of the different retrocue conditions until the cue presentation at 2.4 s following trial onset, encoding of two syllables led to equivalent evoked activity across conditions.

However, the evoked responses of the two retrocue conditions significantly diverged from the onset of the visual retrocue. Compared to the neutral retrocue, the valid retrocue trials exhibited enhanced amplitudes of evoked responses from the visual retrocue presentation and throughout the 2 s retention phase. The multilevel permutation-based statistical test on the ERPs revealed two significant clusters exhibiting a “valid > neutral” effect during the visual cue presentation phase (1, 2.59–2.70 s,  $p = 0.013$ ; 2, 2.73–3.31 s,  $p = 0.002$ ). Since the two clusters were in close temporal proximity, time points and electrodes of these clusters were collapsed and averaged as one cluster. As illustrated in Figure 3B (top), the valid retrocue condition exhibited a stronger positivity compared to the neutral retrocue condition during the presentation of the retrocue.

The same permutation-based test on the ERPs revealed four significant clusters of the reverse effect (i.e., “neutral > valid”) during the retention phase (1, 4.18–4.44 s,  $p = 0.013$ ; 2, 4.47–4.58 s,  $p = 0.039$ ; 3, 4.60–4.81 s,  $p = 0.005$ ; 4, 4.93–5.10 s,  $p = 0.020$ ; note that these negative clusters show up as valid > neutral in the

GFP). These four clusters were also collapsed and averaged as one cluster. During the stimulus-free retention phase, the valid retrocue condition exhibited a significantly enhanced negative potential compared to the neutral condition (Fig. 3B, bottom). The enhanced negativity was broadly distributed, but most pronounced at frontocentral electrodes.

We further examined whether these retrocue-related modulations in the evoked response amplitude were related to the modulations in the behavioral performance of valid and neutral retrocue conditions. From each of these clusters during the retrocue presentation and the retention phases separately, individuals’ amplitude differences in evoked responses ( $\text{GFP}_{\text{Valid}}$  vs  $\text{GFP}_{\text{Neutral}}$ ) were correlated with the difference in the perceptual sensitivity between conditions ( $d'_{\text{Valid}}$  vs  $d'_{\text{Neutral}}$ ). The mean GFP difference ( $\text{GFP}_{\text{Valid}}$  vs  $\text{GFP}_{\text{Neutral}}$ ) during the retrocue presentation phase (2.59–3.31 s) was neither related to the difference in perceptual sensitivity ( $d'$ ; Spearman’s  $\rho = 0.18$ ,  $p = 0.46$ ) nor to the difference in perceptual precision ( $\ln k$ ; Spearman’s  $\rho = 0.15$ ,  $p = 0.54$ ). In addition, the mean GFP difference during retention (4.18–5.10 s) did not predict the perceptual sensitivity



**Figure 4.** Time–frequency representations of valid and neutral retrocue trials during the syllable pitch discrimination task. **A**, Grand average oscillatory power averaged across all EEG participants ( $n = 20$ ) and across all (61) scalp electrodes. **B**, Illustrations of significant time–frequency clusters exhibiting valid versus neutral retrocue contrast. **C**, Relative oscillatory power change of a significant cluster exhibiting a neutral > valid retrocue effect (top) and of a cluster exhibiting a valid > neutral effect (bottom). Oscillatory power of the frequency range and electrodes from the corresponding cluster are averaged for each time point. Highlighted regions (orange and green) indicate the time windows of the significant clusters exhibiting retrocue condition differences from the statistical analysis. Shaded error lines indicate  $\pm 1$  within-subjects SE. **D**, The distributions of  $t$  statistics over frequencies and scalp topographical maps of the significant neutral > valid (top) and valid > neutral clusters (bottom). Highlighted electrodes of scalp topographies belong to the corresponding clusters. Colors in the topographical maps correspond to the color bar in **E**. **E**, Source projections of oscillatory power contrast (valid vs neutral;  $13 \pm 4$  Hz) on a standard MNI template brain. Source-projected  $t$  statistic maps of neutral > valid (top) and valid > neutral (bottom) effects are illustrated. For illustration purpose, a threshold of  $|t_{(19)}| \geq 1.5$  is applied. L and R indicate left and right hemisphere, respectively. The color bar indicates  $t$  values of valid–neutral contrast.

difference between retrocue conditions ( $d'_{\text{Valid}}$  vs  $d'_{\text{Neutral}}$ ; Spearman's  $\rho = 0.29$ ;  $p = 0.21$ ). The relationship of mean GFP difference and perceptual precision difference between conditions ( $\ln k_{\text{Valid}}$  vs  $\ln k_{\text{Neutral}}$ ) amounted to a Spearman's  $\rho$  of 0.39 ( $p = 0.09$ ).

### Alpha power reflects precision in working memory

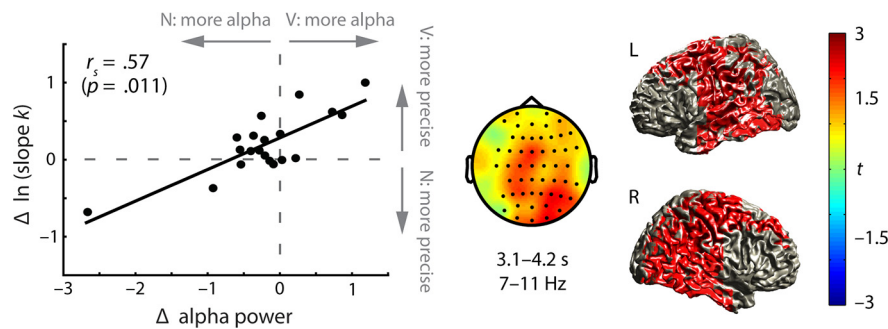
Figure 4A illustrates the dynamics of average oscillatory power across the valid and neutral retrocue trials. As expected for an attention-demanding auditory task, there was a marked enhancement of overall alpha oscillatory power relative to baseline throughout the entire trial period.

The multilevel permutation-based statistical test examining potential retrocue condition contrasts in oscillatory power revealed two significant clusters (Fig. 4B). One cluster was found during the retrocue presentation phase; this cluster exhibited significantly stronger oscillatory power suppression in the valid compared to the neutral retrocue condition (Fig. 4C, top). This suppression was significant not only within alpha range, but also in the broad frequency range (5–40 Hz; 2.7–3.6 s,  $p < 0.001$ ; Fig. 4D, top). Source space result revealed that this power suppression effect was at peak in the left superior parietal cortex ( $t_{(19)} = 3.77$ ; MNI coordinates,  $[-20, -59, 70]$ ; Fig. 4E, top). Nevertheless, the effect was revealed in widely distributed areas including parietal/occipital regions, the bilateral supramarginal gyrus (SMG; BA

40), the right insula (BA 13), the right precentral/postcentral cortex, and the left frontal cortical regions, extending into anterior cingulate cortex (BA 32).

The second significant cluster was found during the stimulus-free retention phase. This cluster showed a significant power enhancement, specifically in alpha and low-beta frequency bands (9–18 Hz) in the valid versus neutral condition (Fig. 4C, bottom); yet, the effect was mostly pronounced within alpha frequency range (4–4.6 s,  $p = 0.038$ ; Fig. 4D, bottom). Source localization showed that this power enhancement emerged primarily from the right superior parietal lobule ( $t_{(19)} = 2.70$ ; MNI coordinates,  $[36, -61, 60]$ ), but extended into bilateral SMG, the left temporal gyrus (BA 21/22), and the medial frontal gyrus (Fig. 4E, bottom).

Next, we examined whether the extent of these oscillatory power modulations predicted the differences in task performance. In a model-based EEG analysis (i.e., a parameter estimate from psychophysical modeling was used as a regressor in a permutation-based statistical test across all time–frequency–electrode bins), we tested whether the extent of retrocue related modulations in overall power across the 1–40 Hz range (valid–neutral) predicted the differences in participants' perceptual precision of syllable pitch discrimination ( $\ln k_{\text{Valid}}$  vs  $\ln k_{\text{Neutral}}$ ). This analysis revealed a significant cluster ( $p = 0.011$ ) only within the alpha frequency range (7–11 Hz) from the later phase of retrocue



**Figure 5.** Correlation of neural and behavioral modulations. Left, Correlation of retrocue-related differences in alpha power and perceptual precision (quantified as the log-transformed slope parameter  $k$  in the psychophysical modeling) between conditions [valid (V) vs neutral (N)]. The topographical map illustrates the correlation effects in the significant correlation cluster (3.1–4.2 s; 7–11 Hz). Right, Source-projected correlation of alpha power and perceptual precision modulations (V–N). For illustration, a threshold of  $|t_{(19)}| \geq 1.5$  is applied. L and R indicate left and right hemisphere, respectively.

presentation (3.1 s) to the middle of the following retention phase (4.2 s). The alpha power differences ( $\alpha_{\text{Valid}}$  vs  $\alpha_{\text{Neutral}}$ ) in this cluster exhibited a significant positive correlation with the individual differences in perceptual precision (Spearman's  $\rho = 0.57$ ;  $p = 0.011$ ; Fig. 5, left). Importantly, this alpha power modulation predicting behavioral perceptual precision modulation reached peak at the right superior/middle frontal gyrus ( $t_{(19)} = 4.69$ ; MNI coordinates, [40, 19, 50]), but was widely distributed into the precentral/postcentral cortical regions, bilateral dorsolateral prefrontal cortex (BA 9), and bilateral temporal cortex (BA 21/22/42; Fig. 5, right).

An analog analysis, using the simpler perceptual sensitivity measure ( $d'_{\text{Valid}}$  vs  $d'_{\text{Neutral}}$ ) as regressor, did not reveal any significant cluster ( $p > 0.18$ ).

## Discussion

Can selective attention to an auditory object, not physically present but only held in memory, improve the representation of this object? If so, what are the neural mechanisms supporting this improvement? Here, we investigated these questions with retrocues that directed attention to task-relevant objects in auditory working memory.

### Selective attention to memory objects enhances task performance and representational precision

Our behavioral results revealed beneficial effects of retrospective attention to a specific syllable in memory. Consistent with previous findings that retrospective attention facilitates auditory (Backer and Alain, 2012; Kumar et al., 2013; Backer et al., 2015) and visual working memory performance (Griffin and Nobre, 2003; Makovski et al., 2008; Sligte et al., 2008; Kuo et al., 2009), a retrocue providing valid information about the upcoming probe led to faster and more accurate responses than uninformative neutral cue trials.

Importantly, the psychophysical modeling results transcended this by revealing how selective attention benefits memory performance, namely, through enhanced precision of the attended syllable's representation. We used a modeling approach to obtain a fine-grained measure of perceptual precision established in the visual literature (Zhang and Luck, 2008; Bays and Husain, 2008; Murray et al., 2013).

Notably, our findings in the auditory domain differ from the findings of a previous visual study that saw no evidence of retrospective attention increasing the precision of the attended objects in memory (Murray et al., 2013). This mismatch notwithstanding,

our behavioral and psychophysical modeling results are consistent with the view that retrospective attention enhances memory representations (Lepsien et al., 2011; Rerko and Oberauer, 2013; Souza et al., 2014). In the following sections, we discuss the neural mechanisms supporting such benefit with retrospective attention.

### Retroactive cues affect neural dynamics of object retention in memory

Our EEG results demonstrate that retrospective attention to objects in auditory working memory is associated with neural modulations of both ERPs and oscillatory power. We suggest that these modulations reflect active engagement of neural resources to maintain the cued items in memory.

As in the study by Backer et al. (2015), we found a greater frontocentral sustained negativity in valid than neutral retrocue trials during the stimulus-free retention phase. This sustained negativity might be a variant of the CNV, indicating anticipation to process an imperative stimulus (Walter et al., 1964; Loveless and Sanford, 1975; Chennu et al., 2013) and the degree of allocating auditory attention (Wöstmann et al., 2015a). Thus, the increased negativity observed here in valid trials may reflect enhanced attention allocation to the cued object in auditory memory.

Our results further suggest that the alpha power dynamics following retrocue onset indicate the benefits of attention, specifically on representational precision of objects in working memory: the extent of individuals' alpha power modulations by conditions predicted the modulations of representational precision of syllable objects in memory. Also, the few individuals who showed enhanced alpha power in neutral trials rather than in valid trials retained precise syllable representations in memory comparable to or even better than the precision of the attended syllables in valid trials. This finding is consistent with the Wilsch et al.'s (2015) account on a compensatory mechanism of alpha power in facilitating performance. Thus, we suggest that an overall increase of alpha power is beneficial, especially for highlighting internal representations of objects in auditory memory.

So what is the mechanism by which alpha power highlights representations of memory objects? According to the "functional inhibition" account of alpha power (Klimesch et al., 2007; Jensen and Mazaheri, 2010; Klimesch, 2012), enhancement of representational precision of memory objects is possibly achieved through inhibition of irrelevant information. During retention, we observed that valid retrocues induced overall enhancement of alpha/low-beta power (Fig. 4C,D). This pattern is consistent with a similar functional role of alpha and beta power on memory processing (Hanslmayr et al., 2012; Waldhauser et al., 2012), and may suggest that increased alpha/beta power helps to suppress irrelevant (i.e., uncued) syllable processing, and thus indirectly supports the maintenance of a relevant (i.e., cued) object in memory.

The source localization results suggest contribution of both domain-general and domain-specific regions to maintenance of the cued object in memory. The valid-cue-related alpha/beta power enhancement during retention was at peak in the posterior/parietal cortex, typical regions implicated as an alpha oscillatory network (Foxe et al., 1998). Furthermore, alpha power modulations that significantly predicted perceptual precision benefits with retrospective attention were localized in broad re-



gions of the frontal cortex, including lateral prefrontal cortex, implicated in maintenance of task-relevant internal representations (Curtis and D'Esposito, 2003). Involvement of these regions is suggestive of functional inhibition being implemented via "top-down" attentional control from frontoparietal and dorsal attentional networks (Fox et al., 2006; Dosenbach et al., 2007; Sadaghiani et al., 2010). Moreover, as observed by Obleser et al. (2012), alpha power enhancement during retention was also localized in the SMG and superior/middle temporal regions related to auditory/verbal processing. The SMG has been implicated in pitch memory (Gaab et al., 2003), verbal working memory (Buchsbaum and D'Esposito, 2008; Obleser and Eisner, 2009), and acoustic change detection of syllables (Celsis et al., 1999; Zevin and McCandliss, 2005; Joanisse et al., 2007). Also, the superior/middle temporal cortical regions are related to auditory and speech perception (Liebenthal et al., 2005; Desai et al., 2008; for review, see Obleser and Eisner, 2009). Overall, these results indicate that domain-general executive attention networks as well as domain-specific regions (Strauß et al., 2014a) contribute to the maintenance of precise syllable object representations.

However, our study cannot disentangle whether the perceptual precision benefit with valid retrocues is due to an enhancement of cued objects, suppression of uncued objects, or both—a persistent ambiguity in alpha-power-based analyses that contrast task-irrelevant and task-relevant demands (Klimesch et al., 2007; Palva et al., 2011).

### What can be inferred on the functional mechanisms of retrospective attention?

Among the various mechanisms of valid retrocues postulated by the visual literature (Souza et al., 2014), one dominant notion is that valid retrocues facilitate performance by removing irrelevant objects from working memory (Oberauer, 2001; Oberauer et al., 2012). The study by Kuo et al. (2012) supported this removal account as valid retrocues reduced the contralateral delay activity (CDA), a neural marker of visual working memory load (Vogel and Machizawa, 2004; Vogel et al., 2005).

However, our ERP and alpha power results exhibit an opposite pattern to the removal account, which would predict valid retrocues to reduce neural signatures of auditory memory. A sustained anterior negativity as seen here is the auditory analog of the CDA; it is a retention-related frontocentral negativity component, the magnitude of which increases with auditory memory load (Guimond et al., 2011; Lefebvre et al., 2013). Thus, our results of larger cue-related magnitudes of this negativity contradict the removal account. Likewise, alpha power indexes working memory load and increases with the number of items held in memory across sensory modalities (Jensen et al., 2002; Obleser et al., 2012), and we observed a pattern incompatible with the removal account, that is, enhanced oscillatory power during retention in valid trials. The present data thus unanimously demonstrate that valid retrocues recruit neural resources to retain and highlight the attended representations, rather than freeing resources by removing irrelevant objects from memory.

The current experiment investigated the retrocuing effect with only two auditory objects; therefore, it is still an open question whether the use of retrocues depends on differences in auditory versus visual modalities (Demany et al., 2010) or the amount of spare capacity in working memory (Matsukura et al., 2007; Astle et al., 2012). For instance, when memory capacity is exceeded, retrospective attention may instead lead to alpha power suppression, thereby supporting the removal account. Future investigations are needed to confirm the different factors that can

impact underlying mechanisms of auditory retrospective attention.

### Neural dynamics reflect cue processing and attentional orientation

During the presentation of retrocues, valid versus neutral cues elicited differential patterns of ERPs and alpha power extending to a wide frequency range. As typically shown by modulation of late positive responses with task-relevance, context updating, and the selection process (Desmedt, 1980; Donchin and Coles, 1988; Polich, 2007), we found that valid retrocues increased the amplitude of positive-going evoked responses during cue presentation (Fig. 3B). Moreover, alpha as well as beta power suppression has been associated with the degree of semantic information (Klimesch, 1997, 1999, 2011; Hanslmayr et al., 2009; Shahin et al., 2009), successful memory encoding of task-relevant information (Hanslmayr et al., 2012), and attention allocation for memory retrieval (Pesonen et al., 2006; Mazaheri et al., 2014; Backer et al., 2015). The source of valid-cue-related alpha/beta power suppression emerged in broad regions including the posterior/parietal cortex and the cingulate and frontal cortices. This pattern suggests less functional inhibition of these general attentional networks, which in turn indicates active engagement of the cortical regions, such as cingulo-opercular and frontoparietal networks (Dosenbach et al., 2007), for processing task-relevant information. These findings are consistent with alpha/beta power suppression observed here for valid (i.e., written-syllable) retrocues. Thus, we suggest that these neural effects during cue presentation not only reflect encoding and/or interpretation of the visually presented retrocue information, but also indicate attentional orientation to the task-relevant object in memory.

### Conclusions

Surprisingly few studies have investigated the role of retrospective attention, particularly in the auditory modality. The present study elucidates the underlying neural mechanisms by which retrospective attention facilitates auditory working memory performance. By using psychophysical modeling and model-based EEG analysis, we demonstrate that selective attention to an auditory object in memory improves representational precision of the attended object, and neural modulations of both ERPs and alpha oscillatory power reflect benefits of top-down attention to specific object representations in memory. In sum, our findings provide evidence that, rather than removing task-irrelevant items from memory, retrospective attention to auditory memory content recruits neural resources to strengthen internal representations of the attended object.

### References

- Astle DE, Summerfield J, Griffin I, Nobre AC (2012) Orienting attention to locations in mental representations. *Atten Percept Psychophys* 74:146–162. [CrossRef Medline](#)
- Backer KC, Alain C (2012) Orienting attention to sound object representations attenuates change deafness. *J Exp Psychol Hum Percept Perform* 38:1554–1566. [CrossRef Medline](#)
- Backer KC, Alain C (2014) Attention to memory: orienting attention to sound object representations. *Psychol Res* 78:439–452. [CrossRef Medline](#)
- Backer KC, Binns MA, Alain C (2015) Neural dynamics underlying attentional orienting to auditory representations in short-term memory. *J Neurosci* 35:1307–1318. [CrossRef Medline](#)
- Bays PM, Husain M (2008) Dynamic shifts of limited working memory resources in human vision. *Science* 321:851–854. [CrossRef Medline](#)
- Buchsbaum BR, D'Esposito M (2008) The search for the phonological store: from loop to convolution. *J Cogn Neurosci* 20:762–778. [CrossRef Medline](#)

- Celsis P, Boulanouar K, Doyon B, Ranjeva JP, Berry I, Nespoulous JL, Chollet F (1999) Differential fMRI responses in the left posterior superior temporal gyrus and left supramarginal gyrus to habituation and change detection in syllables and tones. *Neuroimage* 9:135–144. [CrossRef Medline](#)
- Chennu S, Noreika V, Gueorguiev D, Blenkmann A, Kochen S, Ibáñez A, Owen AM, Bekinschtein TA (2013) Expectation and attention in hierarchical auditory prediction. *J Neurosci* 33:11194–11205. [CrossRef Medline](#)
- Cohen MX (2014) *Analyzing neural time series data: theory and practice*. Cambridge, MA: MIT.
- Curtis CE, D'Esposito M (2003) Persistent activity in the prefrontal cortex during working memory. *Trends Cogn Sci* 7:415–423. [CrossRef Medline](#)
- Debener S, Thorne J, Schneider TR, Viola FC (2010) Using ICA for the analysis of multi-channel EEG data. In: *Simultaneous EEG and fMRI: recording, analysis, and application* (Ullsperger M, Debener S, eds), pp 121–133. New York: Oxford UP.
- Demany L, Semal C, Cazalets JR, Pressnitzer D (2010) Fundamental differences in change detection between vision and audition. *Exp Brain Res* 203:261–270. [CrossRef Medline](#)
- Desai R, Liebenthal E, Waldron E, Binder JR (2008) Left posterior temporal regions are sensitive to auditory categorization. *J Cogn Neurosci* 20:1174–1188. [CrossRef Medline](#)
- Desmedt JE (1980) P300 in serial tasks: An essential post-decision closure mechanism. In: *Progress in Brain Research* (Kornhubek HH, Deecke L, eds), pp 682–686. Amsterdam: Elsevier.
- Donchin E, Coles MGH (1988) Is the P300 component a manifestation of context updating? *Behav Brain Sci* 11:357–374. [CrossRef](#)
- Dosenbach NU, Fair DA, Miezin FM, Cohen AL, Wenger KK, Dosenbach RA, Fox MD, Snyder AZ, Vincent JL, Raichle ME, Schlaggar BL, Petersen SE (2007) Distinct brain networks for adaptive and stable task control in humans. *Proc Natl Acad Sci U S A* 104:11073–11078. [CrossRef Medline](#)
- Fox MD, Snyder AZ, Zacks JM, Raichle ME (2006) Coherent spontaneous activity accounts for trial-to-trial variability in human evoked brain responses. *Nat Neurosci* 9:23–25. [CrossRef Medline](#)
- Foxe JJ, Simpson GV, Ahlfors SP (1998) Parieto-occipital approximately 10 Hz activity reflects anticipatory state of visual attention mechanisms. *Neuroreport* 9:3929–3933. [CrossRef Medline](#)
- Gaab N, Gaser C, Zaehle T, Jancke L, Schlaug G (2003) Functional anatomy of pitch memory—an fMRI study with sparse temporal sampling. *Neuroimage* 19:1417–1426. [CrossRef Medline](#)
- Gazzaley A, Nobre AC (2012) Top-down modulation: bridging selective attention and working memory. *Trends Cogn Sci* 16:129–135. [CrossRef Medline](#)
- Griffin IC, Nobre AC (2003) Orienting attention to locations in internal representations. *J Cogn Neurosci* 15:1176–1194. [CrossRef Medline](#)
- Griffiths TD, Warren JD (2004) What is an auditory object? *Nat Rev Neurosci* 5:887–892. [Medline](#)
- Gross J, Kujala J, Hamalainen M, Timmermann L, Schnitzler A, Salmelin R (2001) Dynamic imaging of coherent sources: Studying neural interactions in the human brain. *Proc Natl Acad Sci U S A* 98:694–699. [CrossRef Medline](#)
- Guimond S, Vachon F, Nolden S, Lefebvre C, Grimault S, Jolicœur P (2011) Electrophysiological correlates of the maintenance of the representation of pitch objects in acoustic short-term memory. *Psychophysiology* 48:1500–1509. [CrossRef Medline](#)
- Haegens S, Osipova D, Oostenveld R, Jensen O (2010) Somatosensory working memory performance in humans depends on both engagement and disengagement of regions in a distributed network. *Hum Brain Mapp* 31:26–35. [Medline](#)
- Hanslmayr S, Spitzer B, Bäuml KH (2009) Brain oscillations dissociate between semantic and nonsemantic encoding of episodic memories. *Cereb Cortex* 19:1631–1640. [CrossRef Medline](#)
- Hanslmayr S, Staudigl T, Fellner MC (2012) Oscillatory power decreases and long-term memory: the information via desynchronization hypothesis. *Front Hum Neurosci* 6:74. [Medline](#)
- Jensen O, Mazaheri A (2010) Shaping functional architecture by oscillatory alpha activity: Gating by inhibition. *Front Hum Neurosci* 4:186. [Medline](#)
- Jensen O, Gelfand J, Kounios J, Lisman JE (2002) Oscillations in the alpha band (9–12 Hz) increase with memory load during retention in a short-term memory task. *Cereb Cortex* 12:877–882. [CrossRef Medline](#)
- Joanisse MF, Zevin JD, McCandliss BD (2007) Brain mechanisms implicated in the preattentive categorization of speech sounds revealed using fMRI and a short-interval habituation trial paradigm. *Cereb Cortex* 17:2084–2093. [CrossRef Medline](#)
- Klimesch W (1997) EEG-alpha rhythms and memory processes. *Int J Psychophysiol* 26:319–340. [CrossRef Medline](#)
- Klimesch W (1999) EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res Rev* 29:169–195. [CrossRef Medline](#)
- Klimesch W (2011) Evoked alpha and early access to the knowledge system: the P1 inhibition timing hypothesis. *Brain Res* 1408:52–71. [CrossRef Medline](#)
- Klimesch W (2012) Alpha-band oscillations, attention, and controlled access to stored information. *Trends Cogn Sci* 16:606–617. [CrossRef Medline](#)
- Klimesch W, Sauseng P, Hanslmayr S (2007) EEG alpha oscillations: The inhibition–timing hypothesis. *Brain Res Rev* 53:63–88. [CrossRef Medline](#)
- Kumar S, Joseph S, Pearson B, Teki S, Fox ZV, Griffiths TD, Husain M (2013) Resource allocation and prioritization in auditory working memory. *Cogn Neurosci* 4:12–20. [CrossRef Medline](#)
- Kuo BC, Rao A, Lepsien J, Nobre AC (2009) Searching for targets within the spatial layout of visual short-term memory. *J Neurosci* 29:8032–8038. [CrossRef Medline](#)
- Kuo BC, Stokes MG, Nobre AC (2012) Attention modulates maintenance of representations in visual short-term memory. *J Cogn Neurosci* 24:51–60. [CrossRef Medline](#)
- Lefebvre C, Vachon F, Grimault S, Thibault J, Guimond S, Peretz I, Zatorre RJ, Jolicœur P (2013) Distinct electrophysiological indices of maintenance in auditory and visual short-term memory. *Neuropsychologia* 51:2939–2952. [CrossRef Medline](#)
- Lehmann D, Skrandies W (1980) Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalogr Clin Neurophysiol* 48:609–621. [CrossRef Medline](#)
- Lepsien J, Thornton I, Nobre AC (2011) Modulation of working-memory maintenance by directed attention. *Neuropsychologia* 49:1569–1577. [CrossRef Medline](#)
- Liebenthal E, Binder JR, Spitzer SM, Possing ET, Medler DA (2005) Neural substrates of phonemic perception. *Cereb Cortex* 15:1621–1631. [CrossRef Medline](#)
- Loveless NE, Sanford AJ (1975) The impact of warning signal intensity on reaction time and components of the contingent negative variation. *Biol Psychol* 2:217–226. [CrossRef Medline](#)
- Macmillan NA, Creelman CD (2004) *Detection theory: a user's guide*, Ed 2. London: Psychology.
- Makovski T, Jiang YV (2007) Distributing versus focusing attention in visual short-term memory. *Psychon Bull Rev* 14:1072–1078. [CrossRef Medline](#)
- Makovski T, Sussman R, Jiang YV (2008) Orienting attention in visual working memory reduces interference from memory probes. *J Exp Psychol Learn Mem Cogn* 34:369–380. [CrossRef Medline](#)
- Manza P, Hau CL, Leung HC (2014) Alpha power gates relevant information during working memory updating. *J Neurosci* 34:5998–6002. [CrossRef Medline](#)
- Matsukura M, Luck SJ, Vecera SP (2007) Attention effects during visual short-term memory maintenance: protection or prioritization? *Percept Psychophys* 69:1422–1434. [CrossRef Medline](#)
- Mazaheri A, van Schouwenburg MR, Dimitrijevic A, Denys D, Cools R, Jensen O (2014) Region-specific modulations in oscillatory alpha activity serve to facilitate processing in the visual and auditory modalities. *Neuroimage* 87:356–362. [CrossRef Medline](#)
- Murray AM, Nobre AC, Clark IA, Cravo AM, Stokes MG (2013) Attention restores discrete items to visual short-term memory. *Psychol Sci* 24:550–556. [CrossRef Medline](#)
- Oberauer K (2001) Removing irrelevant information from working memory: a cognitive aging study with the modified Sternberg task. *J Exp Psychol Learn Mem Cogn* 27:948–957. [CrossRef Medline](#)
- Oberauer K, Lewandowsky S, Farrell S, Jarrold C, Greaves M (2012) Modeling working memory: an interference model of complex span. *Psychon Bull Rev* 19:779–819. [CrossRef Medline](#)
- Obleser J, Eisner F (2009) Pre-lexical abstraction of speech in the auditory cortex. *Trends Cogn Sci* 13:14–19. [CrossRef Medline](#)
- Obleser J, Weisz N (2012) Suppressed Alpha oscillations predict intelligibility of speech and its acoustic details. *Cereb Cortex* 22:2466–2477. [CrossRef Medline](#)
- Obleser J, Wöstmann M, Hellbernd N, Wilsch A, Maess B (2012) Adverse

- listening conditions and memory load drive a common  $\alpha$  oscillatory network. *J Neurosci* 32:12376–12383. [CrossRef Medline](#)
- Oostenveld R, Stegeman DF, Praamstra P, van Oosterom A (2003) Brain symmetry and topographic analysis of lateralized event-related potentials. *Clin Neurophysiol* 114:1194–1202. [CrossRef Medline](#)
- 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. [Medline](#)
- O'Sullivan JA, Shamma SA, Lalor EC (2015) Evidence for neural computations of temporal coherence in an auditory scene and their enhancement during active listening. *J Neurosci* 35:7256–7263. [CrossRef Medline](#)
- Palva S, Kulashekhar S, Hämäläinen M, Palva JM (2011) Localization of cortical phase and amplitude dynamics during visual working memory encoding and retention. *J Neurosci* 31:5013–5025. [CrossRef Medline](#)
- Pertsov Y, Bays PM, Joseph S, Husain M (2013) Rapid forgetting prevented by retrospective attention cues. *J Exp Psychol Hum Percept Perform* 39:1224–1231. [CrossRef Medline](#)
- Pesonen M, Björnberg CH, Hämäläinen H, Krause CM (2006) Brain oscillatory 1–30 Hz EEG ERD/ERS responses during the different stages of an auditory memory search task. *Neurosci Lett* 399:45–50. [CrossRef Medline](#)
- Poch C, Campo P, Barnes GR (2014) Modulation of alpha and gamma oscillations related to retrospectively orienting attention within working memory. *Eur J Neurosci* 40:2399–2405. [CrossRef Medline](#)
- Polich J (2007) Updating P300: an integrative theory of P3a and P3b. *Clin Neurophysiol* 118:2128–2148. [CrossRef Medline](#)
- Rerko L, Oberauer K (2013) Focused, unfocused, and defocused information in working memory. *J Exp Psychol Learn Mem Cogn* 39:1075–1096. [CrossRef Medline](#)
- Sadaghiani S, Scheeringa R, Lehongre K, Morillon B, Giraud AL, Kleinschmidt A (2010) Intrinsic connectivity networks, alpha oscillations, and tonic alertness: a simultaneous electroencephalography/functional magnetic resonance imaging Study. *J Neurosci* 30:10243–10250. [CrossRef Medline](#)
- Shahin AJ, Picton TW, Miller LM (2009) Brain oscillations during semantic evaluation of speech. *Brain Cogn* 70:259–266. [CrossRef Medline](#)
- Shinn-Cunningham BG (2008) Object-based auditory and visual attention. *Trends Cogn Sci* 12:182–186. [CrossRef Medline](#)
- Sligte IG, Scholte HS, Lamme VA (2008) Are there multiple visual short-term memory stores? *PLoS One* 3:e1699. [CrossRef Medline](#)
- Souza AS, Rerko L, Oberauer K (2014) Unloading and reloading working memory: Attending to one item frees capacity. *J Exp Psychol Hum Percept Perform* 40:1237–1256. [CrossRef Medline](#)
- Strauß A, Wöstmann M, Obleser J (2014a) Cortical alpha oscillations as a tool for auditory selective inhibition. *Front Hum Neurosci* 8:350. [Medline](#)
- Strauß A, Kotz SA, Scharinger M, Obleser J (2014b) Alpha and theta brain oscillations index dissociable processes in spoken word recognition. *Neuroimage* 97:387–395. [CrossRef Medline](#)
- Tuladhar AM, ter Huurne NT, Schoffelen JM, Maris E, Oostenveld R, Jensen O (2007) Parieto-occipital sources account for the increase in alpha activity with working memory load. *Hum Brain Mapp* 28:785–792. [CrossRef Medline](#)
- van den Brink RL, Wynn SC, Nieuwenhuis S (2014) Post-error slowing as a consequence of disturbed low-frequency oscillatory phase entrainment. *J Neurosci* 34:11096–11105. [CrossRef Medline](#)
- Vogel EK, Machizawa MG (2004) Neural activity predicts individual differences in visual working memory capacity. *Nature* 428:748–751. [CrossRef Medline](#)
- Vogel EK, McCollough AW, Machizawa MG (2005) Neural measures reveal individual differences in controlling access to working memory. *Nature* 438:500–503. [CrossRef Medline](#)
- Waldhauser GT, Johansson M, Hanslmayr S (2012) Alpha/beta oscillations indicate inhibition of interfering visual memories. *J Neurosci* 32:1953–1961. [CrossRef Medline](#)
- Wallis G, Stokes M, Cousijn H, Woolrich M, Nobre AC (2015) Frontoparietal and cingulo-opercular networks play dissociable roles in control of working memory. *J Cogn Neurosci* 27:2019–2034. [CrossRef Medline](#)
- Walter WG, Cooper R, Aldridge VJ, McCallum WC, Winter AL (1964) Contingent negative variation: an electric sign of sensorimotor association and expectancy in the human brain. *Nature* 203:380–384. [CrossRef Medline](#)
- Weisz N, Hartmann T, Müller N, Lorenz I, Obleser J (2011) Alpha rhythms in audition: cognitive and clinical perspectives. *Front Psychol* 2:73. [Medline](#)
- Wilsch A, Henry MJ, Herrmann B, Maess B, Obleser J (2015) Alpha oscillatory dynamics index temporal expectation benefits in working memory. *Cereb Cortex* 25:1938–1946. [CrossRef Medline](#)
- Wöstmann M, Schröger E, Obleser J (2015a) Acoustic detail guides attention allocation in a selective listening task. *J Cogn Neurosci* 27:988–1000. [CrossRef Medline](#)
- Wöstmann M, Herrmann B, Wilsch A, Obleser J (2015b) Neural alpha dynamics in younger and older listeners reflect acoustic challenges and predictive benefits. *J Neurosci* 35:1458–1467. [CrossRef Medline](#)
- Zevin JD, McCandliss BD (2005) Dishabituation of the BOLD response to speech sounds. *Behav Brain Funct* 1:4. [CrossRef Medline](#)
- Zhang W, Luck SJ (2008) Discrete fixed-resolution representations in visual working memory. *Nature* 453:233–235. [CrossRef Medline](#)