Neural dynamics supporting auditory long-term memory effects on target detection[☆]

Jacqueline Zimmermann [a,b], Bernhard Ross [a,c,d], Morris Moscovitch [a,b], Claude Alain [a,b,d,e,*]

[a] Rotman Research Institute, Psychology, University of Toronto, Ontario, Canada
[b] Department of Psychology, University of Toronto, Ontario, Canada
[c] Department of Medical Biophysics, University of Toronto, Ontario, Canada
[d] Institute of Medical Sciences, University of Toronto, Ontario, Canada
[e] Faculty of Music, University of Toronto, Ontario, Canada

A B S T R A C T

Auditory long-term memory has been shown to facilitate signal detection. However, the nature and timing of the cognitive processes supporting such benefits remain equivocal. We measured neuroelectric brain activity while young adults were presented with a contextual memory cue designed to assist with the detection of a faint pure tone target embedded in an audio clip of an everyday environmental scene (e.g., the soundtrack of a restaurant). During an initial familiarization task, participants heard such audio clips, half of which included a target sound (memory cue trials) at a specific time and location (left or right ear), as well as audio clips without a target (neutral trials). Following a 1-h or 24-h retention interval, the same audio clips were presented, but now all included a target. Participants were asked to press a button as soon as they heard the pure tone target. Overall, participants were faster and more accurate during memory than neutral cue trials. The auditory contextual memory effects on performance coincided with three temporally and spatially distinct neural modulations, which encompassed changes in the amplitude of event-related potential as well as changes in theta, alpha, beta and gamma power. Brain electrical source analyses revealed greater source activity in memory than neutral cue trials in the right superior temporal gyrus and left parietal cortex. Conversely, neutral trials were associated with greater source activity than memory cue trials in the left posterior medial temporal lobe. Target detection was associated with increased negativity (N2), and a late positive (P3b) wave at frontal and parietal sites, respectively. The effect of auditory contextual memory on brain activity preceding target onset showed little lateralization. Together, these results are consistent with contextual memory facilitating retrieval of target-context associations and deployment and management of auditory attentional resources to when the target occurred. The results also suggest that the auditory cortices, parietal cortex, and medial temporal lobe may be parts of a neural network enabling memory-guided attention during auditory scene analysis.

Expectations and long-term knowledge have a strong influence on how visual and auditory attention may be allocated (Zimmermann et al., 2016). For instance, observers are faster at detecting a target stimulus if it occurs within a familiar stimulus array (Chun and Jiang, 1998) or photograph (Patai et al., 2012; Stokes et al., 2012; Summerfield et al., 2006). In visual studies, this effect of contextual memory on behavioural performance can last up to 24 h after learning (Patai et al., 2012; Summerfield et al., 2011), and has been taken as evidence that long-term memory biases the deployment of visual-spatial attention (Patai et al.,

2012; Stokes et al., 2012; Summerfield et al., 2006, 2011). In the auditory domain, familiar voices (Johnsrude et al., 2013), or a self-referential stimulus such as hearing one's own name, can capture attention (Bey and McAdams, 2002; Cherry, 1953; Dowling et al., 1987). Zimmerman, Moscovitch and Alain (2017) also showed in a controlled lab setting that acquired auditory long-term memory (LTM) can facilitate task performance. In an initial learning phase, participants learned target-to-context associations by pairing an audio clip (e.g., sounds in a busy restaurant) with a lateralized pure tone (i.e., target). Participants were also presented

with audio clips that did not include a target (neutral condition). While both memory and neutral trials became equally familiar during the learning task, only memory trials were paired with a target, allowing a context (audio clip) to target association. Following a 1-h retention interval, the same audio clips were presented, but now all included a target. Participants were more accurate and faster in the memory than in the neutral condition, which suggests that learned context-to-target associations mediate auditory attentional resources.

One important question concerns the nature and timing of the cognitive processes that underlie the effects of auditory LTM on target detection. For instance, learned context-to-target associations are likely to trigger recollection-like processes that engage episodic memory (e.g., sound in a busy restaurant and a left-lateralized target). Participant's recollections could then be used to steer attention toward the location of the anticipated target. Prior visual memory-guided attention studies have reported changes in brain activity that are consistent with recollection processes and deployment of spatial attention (Patai et al., 2012; Stokes, 2011; Summerfield et al., 2011). In these studies, participants were presented with a picture that could have been associated with the target (memory cue scene) or not (neutral cue scene). After a delay, the same picture (probe) was presented with the target. Memory and neutral cue scenes generated distinct event-related potential (ERP) signatures, with memory-cue scenes showing a more negative amplitude contralateral to the anticipated target, and are thought to index anticipatory attention (Summerfield et al., 2011). For the probe scene, ERPs revealed modulation of target selection, marked by a negative modulation posterior and contralateral to the target (i.e., N2pc component), which has been associated with selection of task-relevant information and suppression of distracters (Patai et al., 2012; Summerfield et al., 2011).

Prior research examining auditory spatial attention has also revealed a negative modulation analogous to the N2pc, albeit with a more anterior distribution (N2ac) (Gamble and Luck, 2011; Gamble and Woldorff, 2015; Lewald et al., 2016). Moreover, studies focusing on brain oscillations have shown lateralization of alpha power in response to the target location (Klatt et al., 2018, 2020; Müller and Weisz, 2012), which are more prominent over parietal and parieto-occipital areas between 400 and 1000 ms post-stimulus. Studies examining the interplay between attention and auditory memory have also observed changes in theta (Del Giudice et al., 2016; del Giudice et al., 2014; Lim et al., 2015), beta (Backer et al., 2015; Ishii et al., 2009; Lim et al., 2015), and gamma (Gurtubay et al., 2004; Ishii et al., 2009; Kaiser et al., 2008; Lim et al., 2015) power, but it is unclear whether memory-guided attention is also associated with changes in these frequency bands. Both the N2ac and alpha power lateralization are thought to index an orientation to the sound location and could be used to examine whether contextual memory biases the deployment of attention to a lateralized auditory target. If the behavioural benefits associated with auditory contextual memory are related to the deployment of auditory spatial attention, then ERPs indexing memory-guided attention should be comprised of modulations analogous to N2ac and alpha-power lateralization.

Contextual memory could also improve task performance by allowing participants to select and prepare their response prior to target presentation, without necessarily engaging auditory spatial attention. That is, the participant's recollection of where the target will occur could allow them to select the appropriate response (e.g., left or right target) and then await the target presentation to execute the response. Prior research has shown that knowing in advance when the auditory target would likely be presented (i.e., temporal expectations) improves its detection in a stream of distracters (Rimmele et al., 2011; Shen and Alain, 2011, 2012; Shen et al., 2016). The latter has been associated with ERP modulations over central scalp areas thought to index anticipatory attention to a specific point in *time*. Thus, if contextual memory mediates temporal expectation, then ERP modulations indexing memory-guided auditory attention should have a more central distribution and should be little affected by the location of the anticipated target.

1. Impact of retention delay on memory-guided attention

Another aim of the present study was to assess the impact of a retention delay on memory-guided attention to auditory stimuli. Given that neural representations of memories are dynamic over time (Winocur and Moscovitch, 2011), memory-guided orienting may rely on different processes and mechanisms as time passes (Buzsaki, 1989, 1998; Drosopoulos et al., 2005; McClelland et al., 1995; Orban et al., 2006; Wagner et al., 2010). For example, attention facilitated by short-term pattern expectations relies on the frontal-striatal-cerebellar brain network, while contextual cueing relies on the medial temporal lobe (Negash et al., 2007b). Theoretically, the hippocampus should come into play when current events are predicted by contextual knowledge evoked by situations and events. This may change, however, with retention intervals longer than several minutes or a few hours of learning, which could affect memory-guided attention. For example, as memory becomes less accessible, physical details are lost, and general contextual information and gist becomes more important (Sekeres et al., 2016). The studies of memory for semantically relevant audio clips have tested recognition or recall only shortly after learning (Cohen et al., 2009; Crutcher and Beer, 2011; Snyder and Gregg, 2011). Hence, it remains unclear whether auditory LTM acquired in a controlled lab setting would also facilitate attention beyond the 1-h interval. In the visual domain, memory-guided attention has been studied after 1 h, a day, or a week of delay (Chun and Jiang, 2003; Patai et al., 2012; Stokes et al., 2012; Summerfield et al., 2011), but the effects of the delay interval have not been compared directly. Moreover, the neural correlates of memory-guided attention to sound objects have yet to be identified. The current study examined whether our ability to use auditory LTM to guide attention persists 24 h after learning.

2. The current study

We used the paradigm developed by Zimmerman et al. (2017). Participants were first presented with audio clips, half of which were paired with a target sound (memory cue trials) at a specific time and location (left or right ear). Following a 1 h or 24-h retention interval, participants were presented with the same audio clips, but now all included a target. On each trial, the audio clip was repeated twice. The first presentation served as a cue and did not include the target (cue audio clip) while the second presentation (i.e., probe audio clip) was comprised of a lateralized soft pure tone target embedded in the audio clip at a specific time. Participants were asked to press a button as quickly as possible when they heard the pure tone target. Our predictions were that response time would be faster for memory than neutral cue trials. We also anticipated a difference in brain activity when the target would be preceded by a memory cue rather than a neutral cue. If memory cues mediate auditory spatial attention, then we should observe a difference in neural activity when the memory cue is associated with a left- or right-lateralized target. However, if contextual memory facilitates task performance by mediating temporal expectations, then neural activity preceding the lateralized target should be comparable.

3. Materials and methods

3.1. Participants

Forty-eight young adults with normal hearing and normal or corrected-to-normal vision were recruited from the research volunteer database at the Rotman Research Institute. One participant was excluded due to technical difficulties during EEG recording, and five participants were excluded due to inadequate explicit memory results.[1] Therefore, a

---

[1] The criterion for rejection was eight or fewer target present trials that were correctly recalled in the explicit memory task.

total of 42 participants were included in the study ($M = 22.5$ years; range 18–30 years; 18 males), with 20 participants in the 1-h retention group ($M = 22.3$ years; range 18–30 years; 8 males) and 22 participants in the 24-h retention group ($M = 22.9$ years; range 19–30 years; 10 males).

Pure tone hearing thresholds were assessed at the octave frequencies between 250 Hz and 8000 Hz. The criterion for normal hearing required thresholds lower than or equal to 25 dB Hearing Level (dB HL) and no more than 15 dB difference between the ears at any test frequency. All participants were right-handed, fluent in English to ensure understanding of the experimental process, and had no history of psychiatric, neurological, or other major illnesses. Participants received monetary compensation for their participation. The study protocol was approved by the Baycrest Research Ethics Board, and the study was carried out in accordance with their recommendations. All participants gave written informed consent in accordance with the Declaration of Helsinki.

3.2. Stimuli

One hundred and four audio clips were retrieved from a public domain sound archive. The clips were stereo sounds chosen to maintain considerable semantic relevance to increase the likelihood that an appropriate association could be formed and labelled in LTM. The same audio clips were used in the learning task, the explicit memory task, and the memory-guided attention task of the study. The clips were edited to the duration of 2500 ms with a 100 ms rise and fall time and were re-sampled to the sampling rate of 44100 Hz. All stimuli were presented through insert earphones (EARTONE 3a) at a presentation level of 60 dB sound pressure level (SPL) on average across stimuli, with some sounds peaking at about 80 dB SPL. The auditory target, which was embedded within the audio clips, was a 500 Hz pure tone with 200 ms duration (20 ms rise and fall time). Acoustic stimuli and visual cues were presented using Presentation software (version 13, Neurobehavioral Systems, Albany, CA).

3.3. Procedure

The participants completed four experimental tasks. The first was a calibration task, which aimed to identify the volume at which the target should be presented in the subsequent tasks. The second task was the learning task in which the participants were presented with audio clips, with half of them, including a left or right-lateralized pure tone target. The third task was the explicit memory task, which was used to identify whether these associations were retained in memory. The fourth was the memory-guided attention task, which measured the participant's ability to use the target-context associations to bias attention towards the target.

3.3.1. Calibration task

For each participant, the volume of the target presented in subsequent tasks was determined using a two-interval, two-alternative forced-choice procedure with a three-down and one-up rule that estimates the 79% correct point on the psychometric function (Levitt, 1971). On each trial, participants were presented with the same audio clip twice (each audio clip was 500 ms in duration) separated by a 500 ms silent interval. The audio clip was selected from a set of four audio clips (i.e., birds chirping, sandpaper, people murmuring, industrial machine). None of these audio clips were used in the subsequent phases of the study. A pure tone target (500 Hz, 500 ms in duration, 50 ms rise/fall time) was embedded in one of the two audio clip presentations. Participants were asked to indicate, by pressing a button, which of the two audio clips included the target. At the beginning of the test, the target intensity was set at 60 dB SPL. The target SPL was calculated by taking an average of the last eight reversals. It was then increased by 10% in the following tasks since many of the soundtracks masked the target when the non-corrected target SPL was used in a pilot study. The participant target volume identified in the calibration task was kept constant throughout the experiment.

3.3.2. Learning task

Half of the 104 audio clips were used for memory trials and the other half for neutral trials. In the memory trials, the audio clip was paired with a pure tone target in the left or right ear. For the neutral trials, there was no pairing between the audio clip and a pure tone. All participants underwent four learning blocks, each block with one presentation of each of the 104 audio clips. The four repetitions of each audio clip were to promote a strong association between the audio clips and the location of the target. The neutral trials were included in the learning phase to yield comparable level of familiarity for the neutral and the memory trials. That is, a memory trace has possibly been formed for the neutral trials, that being a spatially neutral memory trace. The order of memory and neutral trials was random within each block. The audio clips assigned to neutral or memory trials were counterbalanced across participants and groups.

Within the memory trials, the target tone always occurred at 2000 ms after the onset of the audio clip. Participants were instructed to listen for the location of the target within each audio clip and memorize it. They pressed the left, right or down arrow key on a computer keyboard when the target was played from the left side, right side, or if no target was present, respectively. Participants were asked to respond as quickly and as accurately as possible. Participants were given 2000 ms to respond following the offset of the audio clip, and subsequently received visual feedback for 500 ms.

3.3.3. Explicit memory task

Immediately following the learning task, both groups completed a cued recall memory task, which aimed to determine whether participants formed explicit associations between the audio clips and the target location. Participants were presented with the same audio clips as in the learning phase but without the target tone. Participants were given as much time as needed to indicate whether the target had been presented from the left or right ear or whether no target had been present. Subsequently, they rated their confidence in the response using a 4-step scale keypress response, coding "I do not know" responses as 0, "not very confident" responses as 1, "fairly confident" responses as 2, and "very confident" responses as 3.

3.3.4. Memory-guided attention task

Following a 1 h ($N = 20$) or 24 h ($N = 22$) retention interval, participants were presented with the same 104 audio clips from the learning task (including those not correctly recalled at the explicit memory task), each repeated twice to ensure that individuals had sufficient time to access the learned target-context associations (i.e., to ensure sufficient cueing). This design is analogous to prior studies assessing memory-guided attention in the auditory (Zimmermann et al., 2017) and visual domain (Summerfield et al., 2011). A 1000 ms inter-stimulus interval separated the first (cue) audio clip and second (probe) audio clip presentation. The cue audio clip did not include a target tone and served as a retrieval cue to guide attention towards the remembered target location. The probe audio clip always comprised an embedded pure tone target.

For memory trials, the target was always presented at the learned location. For neutral trials, a target was presented to the right or left ear randomly. On each trial, participants pressed the left or right keyboard button as quickly and as accurately as possible when they heard the target. The next trial was presented after the participants' button press. Participants performed the memory-guided attention twice for a grand total of 208 trials. For the neutral trials, the target was presented at the same location in both blocks of trials.

3.4. Behavioural data analysis

For all tasks, trials with reaction times (RTs) shorter than 100 ms or longer than the mean plus twice the standard deviation were excluded as outliers. RTs and %-correct detection were analyzed separately with repeated-measures ANOVA.

For the learning task, the within-subject factors were 'learning block' (four levels, blocks 1–4) and 'trial type' (i.e., memory, neutral). For the explicit memory task, the accuracy of recall for the location of the target within each audio clip was compared to chance level using a one-sample *t*-test. Given three possible response options (left, right, no target) and that all keys could be pressed with the same likelihood, the 33% correct response rate was considered the chance level. Explicit memory was analyzed for all trials to examine memory for target presence and target location, as well as for target-present trials only to examine memory for target location. Further repeated measures ANOVAs assessed the explicit memory for the location of the target tone across reported confidence rating. A significant interaction of confidence (I do not know, low, medium, or high) and correctness (correct response, incorrect response) could indicate participants' awareness of their memory recall.

For the memory-guided attention task, a 2 x 2 mixed-design ANOVA assessed the memory-guided attention across the two retention groups, with memory cue as a within-subject variable (memory vs. neutral cue trials) and retention delay (1 h, one day) as a between-subjects variable. Differences in RTs to detect targets preceded by memory or neutral cues were used to gauge the magnitude of memory-guided attention (Patai et al., 2012; Summerfield et al., 2011; Zimmermann et al., 2017). Brain-behaviour relationships were examined using Pearson correlations. We used the Benjamini-Hochberg method to adjust the familywise p-value for multiple comparisons with q = 0.1, m = 21 (i.e., total number of p values) and p = 0.05 (Hochberg and Benjamini, 1990).

3.5. EEG recording and analysis

The EEG was recorded continuously during the memory-guided attention task using a 76-channel acquisition system (BioSemi Active Two, Amsterdam, The Netherlands). Sixty-six EEG electrodes were positioned on the scalp using a BioSemi head cap, according to the standard 10/20 system, with a Common Mode Sense (CMS) active electrode and Driven Right Leg (DRN) ground electrode. Ten additional electrodes were placed below the hairline (both mastoid, both pre-auricular points, outer canthus of each eye, the inferior orbit of each eye, two facial electrodes) to monitor eye movements and to cover the whole scalp evenly. The neuroelectric activity was DC-100 Hz bandpass filtered, digitized at a rate of 512 Hz, and stored for offline analysis. The memory-guided attention task was administered twice to increase the number of trials in each experimental condition and statistical power for the EEG data analysis.

3.5.1. EEG preprocessing

EEG preprocessing was performed using Brain Electrical Source Analysis Research software (BESA, version 7, MEGIS GmbH, Gräfelfing, Germany). The EEG data were visually inspected to identify segments contaminated by defective electrode(s). No more than eight electrodes were interpolated using values from the surrounding electrodes. The EEG was then re-referenced to the average of all electrodes. The continuous EEG was digitally filtered with a 0.1 Hz high-pass filter (forward, 6dB/octave) and 20 Hz low-pass filter (zero phase, 24 dB/octave). For visualization of an individual trial, we averaged epochs that comprised 500 ms prior to and 7000 ms after the onset of the cue audio clip. For statistical analyses, the data were parsed into three sets of epochs to capture cue-probe-, and target-related activity, respectively. Audio clip epochs were locked to the memory cue or probe onset and were defined as −200 ms prior to audio clip onset to 2000 ms. For the target-related analysis, the continuous EEG was filtered with 0.5 Hz high-pass filter (forward, 6dB/octave) and 30 Hz low-pass filter (zero phase, 24 dB/octave) to better capture the transient acoustic change complex and the P3b response. The target epochs started 200 ms before and ended 800 ms after target onset.

For each participant, a set of ocular movements was identified from the continuous EEG recording and then used to generate spatial components that best account for eye movements. The spatial topographies

were then subtracted from the continuous EEG to correct for lateral and vertical eye movements as well as for eye-blinks. After correcting for eye movements, all experimental files for each participant were then scanned for artifacts; epochs including deflections exceeding 120 μV were marked and excluded from the analysis. The remaining epochs were averaged according to electrode position and experimental conditions. Each average was baseline-corrected with respect to a 200 ms pre-stimulus baseline interval. Only trials where participants correctly detected the target lateralization in the memory trials and correctly rejected a response in the neutral trials were included in the ERP analyses. The average number of trials included in the memory and neutral conditions was 74 (std = 15) and 72 (std = 12), respectively.

3.5.2. Distributed source analysis

For modelling the neural sources of ERPs elicited by the memory and neutral cues, we used an iterative application of Low-Resolution Electromagnetic Tomography (LORETA), which reduces the source space in each iteration. This imaging approach, termed Classical LORETA Analysis Recursively Applied (CLARA), provides more focal localizations of the brain activity and can separate sources located in close vicinity (Beniczky et al., 2016; Dimitrijevic et al., 2013). We used CLARA (BESA version 6.1) with a voxel size of 7 mm in the Talairach space; we found that this default setting is appropriate for the distributed images in most situations (Alain et al., 2017; Shen et al., 2018). The regularization parameters that account for the noise in the data were set with a singular value decomposition cutoff at 0.01%. We used a four shell ellipsoidal head model with a head radius of 85 mm, and thickness for scalp, bone and cerebrospinal fluid of 6, 7, 1 mm, respectively. The relative conductivities were 0.33, 0.33, 0.0042, and 1 for brain, scalp, bone, and cerebrospinal fluid, respectively.

3.5.3. Time-frequency analysis

The time-frequency analysis of the EEG signal power was performed with BESA Research software. The continuous EEG data were first digitally filtered with 1 Hz high-pass (forward, 6dB/octave). Procedures for eye correction and artifact rejection were identical to the ERP analyses. The analysis epoch for the time-frequency analysis consisted of 1000 ms of pre-stimulus activity and 2000 ms of post-stimulus activity time-locked to the onset of the cue or probe audio clips. A complex demodulation method with 1 Hz wide frequency bins and 50 ms time resolution in the range of 2 and 50 Hz was used for decomposing the single-trial EEG data into a time-frequency representation. For eliminating the influence of ERPs, we analyzed brain oscillations after removing the averaged evoked responses. We focused the time-frequency analyses on theta (3–7 Hz), alpha (8–12 Hz), beta (13–29 Hz) and gamma (30–50 Hz) bands. In particular, the alpha and beta bands have been related to auditory reflective attention (Alain et al., 2018; Backer et al., 2015; Lim et al., 2015) and visual memory-guided attention (Stokes, 2011; Summerfield et al., 2011).

The results from the time domain (i.e., ERPs), distributed source analysis, and time-frequency analyses were then exported into BESA Statistics 2.0 for statistical analyses. BESA Statistics 2.0 software automatically identifies clusters in time, frequency, and space using a series of t-tests that compared the time-frequency data between experimental conditions at every time point. This preliminary step identified clusters both in time (adjacent time points) and space (adjacent electrodes) where the ERPs differed between the conditions. For cluster building, we used 4 cm spacing between the electrodes, which led to around four neighbours per channel. We used a cluster alpha of .05 for cluster building. A Monte-Carlo resampling technique (Maris and Oostenveld, 2007) was then used to identify those clusters that had higher values than 95% of all clusters derived by random permutation of the data. This non-parametric permutation statistic is no longer subject to the multiple comparisons problem (for an in-depth overview of permutation statistics as implemented in BESA Statistics see Maris and Oostenveld, 2007). The number of permutations was set at 1000.

4. Results

4.1. Behavioural results

4.1.1. Learning phase

Fig. 1 shows group mean accuracy and RTs for memory and neutral trials as a function of learning blocks. For the accuracy data, the ANOVA yielded a main effect of block (F(3,123) = 12.58, p < 0.001. η^2 = 0.24), trial type (F(1,41) = 117.42, p < 0.001, η^2 = 0.74), and a significant interaction between block and trial type (F(3,123) = 3.06, p = 0.031, η^2 = 0.07). The latter indicates that learning the association between an audio clip and the presence of a target benefitted more from exposure than learning the association between an audio clip and the absence of the target. For the RT data, the ANOVA also yielded a main effect of block (F(3,123) = 16.74, p < 0.001. η^2 = 0.29), trial type (F(1,41) = 220.06, p < 0.001, η^2 = 0.84), and a significant interaction between block and trial type (F(3,123) = 25.17, p < 0.001, η^2 = 0.38). The effects of learning block was greater for neutral than memory trials. Together, these results suggest that participants did create an association between the audio clip and the presence/absence of a pure tone target.

4.1.2. Explicit memory task

In the explicit memory task, participants correctly recalled the target location or the absence of a target with an average of 56%, which was significantly higher than chance (t(41) = 9.21, p < 0.001). Participants from the 1-h retention interval group did better at the explicit memory task than those in the 24-h retention interval group (1 h: 62% ± 16%; 24 h: 51% ± 15%; t(40) = 2.26, p = 0.029), even though both were tested immediately after the learning phase. Confidence ratings for explicit memory correctness was higher for correct than incorrect trials (F(1, 40) = 9.57, p = 0.004) in both groups. There was no difference in confidence rating between the groups (F(1, 40) = 1.10, p = 0.30).

4.1.3. Memory-guided attention task-detection: short and long retention intervals

Fig. 2 shows the short and long delay group mean accuracy and RT as a function of cue type. The main effect of cue on accuracy was not significant (F(1,40) = 1.46, p = 0.234), nor was the main effect of group (F < 1) or the cue by group interaction (F < 1). The ANOVA for RT revealed an effect of cue type (F(1,40) = 12.01, p < 0.001), with faster RTs when the target was preceded by a memory than a neutral cue. The main effect of group was not significant (F < 1) nor was the group × cue interaction (F < 1).

4.2. Electrophysiological results

Fig. 3 shows the group mean ERPs time-locked on the cue onset and the probe onset, as well as the ERPs time-locked on the target onset. The onsets of both audio clips were associated with large transient evoked responses. The target sounds generated an acoustic change complex (ACC) followed by a late positive complex that was maximum at parietal scalp regions.

4.2.1. Responses to the cue audio clip

The sound onsets of both memory and neutral cues generated responses that were largest at fronto-central sites and showed a polarity reversal at inferior temporal-parietal and parietal-occipital sites. This amplitude distribution is consistent with tangential source(s) in temporal regions oriented toward the midline fronto-central scalp area in superior temporal gyrus (Picton et al., 1999; Richer et al., 1989). The transient responses were followed by a sustained potential that was largest over central-parietal areas (Fig. 4).

A cluster analysis of the difference wave between ERPs elicited by memory and neutral cues did not show a significant difference between long and short retention intervals (p > .40), indicating that the effects of cue type on ERP amplitude were little affected by the retention interval. Hence, the data from both retention intervals were combined in subsequent analyses to assess contextual memory effects with a total sample size of 42 participants.

The cluster-level statistic comparing the ERP time series following the cue audio clip between memory and neutral conditions revealed six spatial-temporal clusters. The cluster with the largest difference between memory and neutral cue conditions was located over the frontal scalp area at 723–1678 ms after the cue onset (Fig. 4, Table 1), with larger sustained negativity for the neutral than the memory cue. A second cluster over parietal-occipital areas overlapped in time with the first cluster and showed a more positive response for the neutral than the memory cue. The third cluster over the right frontal areas followed the first two clusters in time. The fourth, fifth and sixth clusters preceded the previous ones and revealed less pronounced modulations at central and right central-parietal areas. To test whether an amplitude difference was associated with the target location, we contrasted ERPs elicited by the cue audio clips when they were paired with a left auditory target versus those paired with a right auditory target. This contrast did not reach statistical significance (p = 0.097).

For each participant and each condition, we modeled scalp-recorded ERPs at each time point using distributed source modeling (i.e., CLARA). Then, we compared the mean source activity for a 60 ms interval centered on the peak ERP difference shown in Fig. 4a The contrast between source activities in the memory and the neutral conditions during the presentation of the cue audio clip yielded a greater source activity in the right auditory cortices for the memory cue than for the neutral cue during the 305–365 ms interval (p = 0.021). For the 505–565 ms interval, the neutral cue was associated with greater source activity in the left posterior medial temporal lobe than the memory cue (p = 0.016). For the 1050–1110 ms interval, the contrast between memory and neutral cue conditions yielded significant differences in source activity in the left parietal and left medial temporal lobes, with the memory cue being associated with greater activity in the parietal cortex than the neutral cue (p = 0.006), while the neutral cue was associated with greater activity in
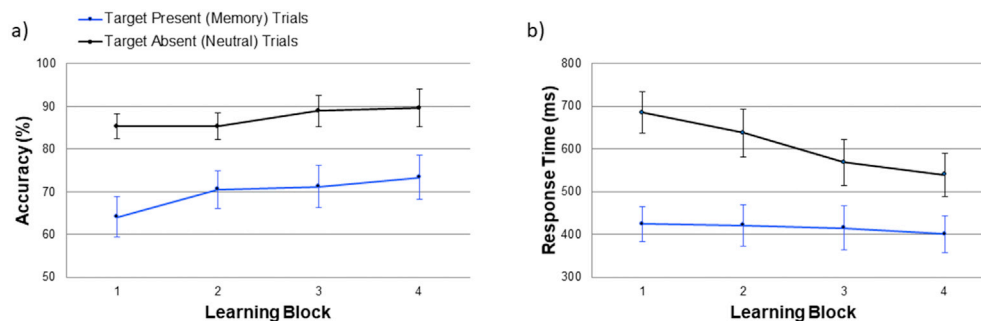

Fig. 1. a) Group mean accuracy as a function of learning block. b) Group mean response time as a function of learning block. The error bars indicate 95% confidence intervals.
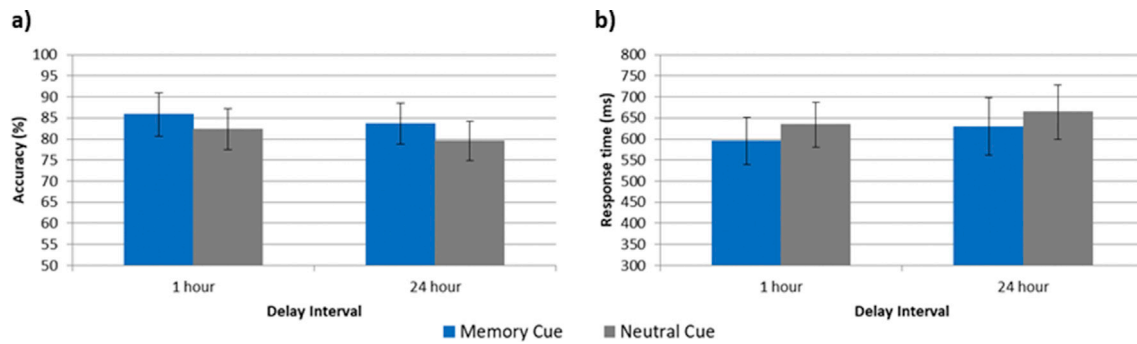
Fig. 2. a) Group mean accuracy as a function of cue type and retention interval for target detection. b) Group mean response time as a function of cue type and retention interval for target detection. The error bars indicate 95% confidence intervals.

medial temporal lobe than the memory cue ($p = 0.03$). Correlations between source activity and RT were computed for each cluster. Although in the memory cue condition, the source activity in the right auditory cortex and left parietal cortex tended to be stronger for faster RTs (r = -0.303 and -0.270, respectively), these correlations were not statistically significant after correcting for multiple comparisons. There was no significant correlation between RTs and source strength for the neutral condition (r = -0.108 and -0.118, ns), nor was the strength of source activity in the other clusters correlated with RTs.

To sum up, contextual auditory memory was associated with three temporally and spatially distinct ERP modulations that began as early as 250 ms after the cue onset. These auditory contextual memory effects were associated with changes in source activity in auditory and parietal cortices. These modulations were little affected by the lateralization of the target and may reflect recollection, anticipation and preparation of participant's response to the expected target.

We also performed time-frequency analyses, which could be more sensitive to time-varying signal changes than the time-domain analyses. Both time-domain and time-frequency analyses complement each other and further strengthen the conclusion regarding the psychological and brain mechanisms associated with memory-guided attention. For this study, we compared oscillatory activity in the theta, alpha, beta and gamma frequency bands during the presentation of the cue with a baseline interval prior to the cue onset. The summary of the clustered-based permutation statistics is presented in Table 2. The theta power

was not statistically different between the cue types ($p > 0.140$). Alpha and beta power decreased in certain time intervals, and gamma power increased during the cue presentation (Fig. 5). Stronger alpha and beta power decreases were observed during the memory cue audio clips compared to neutral cues, while the gamma power increase was stronger during the neutral cues (Table 2).

As for the ERP analysis, we also tested whether the oscillatory activity would differ when audio clips were paired with a left versus a right-lateralized target. For theta power, this contrast yielded one statistically significant cluster ($p = 0.014$), with greater theta power when the cue audio clip was associated with a left rather than a right target (Fig. 6). The difference in theta power was largest at about 600 ms after cue onset over right frontal, central and parietal areas (FP2, AF8, F6, F8, F10, FC6, C6, P6) and left parietal scalp area (CP3, P5, P3, P1, Pz, PO3). There was no significant difference in alpha ($p > 0.17$), beta ($p > 0.57$) or gamma power ($p > 0.16$) between left and right-lateralized targets.

4.2.2. Probe audio clip

Fig. 7 shows ERPs time-locked on the onset of the probe audio clips, 1000 ms after the offset of the cue audio clip, for the memory and neutral cue conditions. The ERPs comprised a transient N1 and P2 wave generated by the onset for the audio clip, which was followed by a small sustained potential at central sites. The target tone embedded in the audio clip generated an ACC at frontal and central sites, which was followed by a large P3b wave at parietal sites. We first tested for differences in ERP amplitude before target onset (i.e., 0–2000 ms). This contrast revealed a small cluster over the mid frontal and parietal areas ($p = 0.029$, 712–920 ms; FP2, AF4, F2, F4, F6, F8, FC4), with the probe following a neutral cue generating a larger amplitude than memory cue. This modulation was unforeseen and may reflect differences in expectations. In the neutral cue condition, participants can only rely on timing information to detect the target, whereas, in the memory trials, participants could use both spatial and timing information. As for the cue audio clips, there was no difference in ERP amplitude when the audio clip was associated with either a left or right auditory target ($p \geq 0.20$). There were no significant differences in source activity between the memory and neutral cue conditions ($p = 0.39$), and no significant spectral power differences were observed between memory and neutral cues. Also, spectral power changes were not statistically different between right and left targets ($p > 0.10$ in all cases).

4.2.3. Target-related activity

The effects of cueing on processing the auditory target were examined after referencing the ERPs to the 200 ms baseline interval before target onset (Fig. 7). The ERPs elicited by the target tones were averaged over the ear of stimulation with the electrodes transposed so that those over the right hemisphere are contralateral to the target location, and those over the left hemisphere are ipsilateral to the target location. Cluster-permutation tests revealed two electrode clusters with differences between memory and neutral cue condition (Table 3). The most prominent
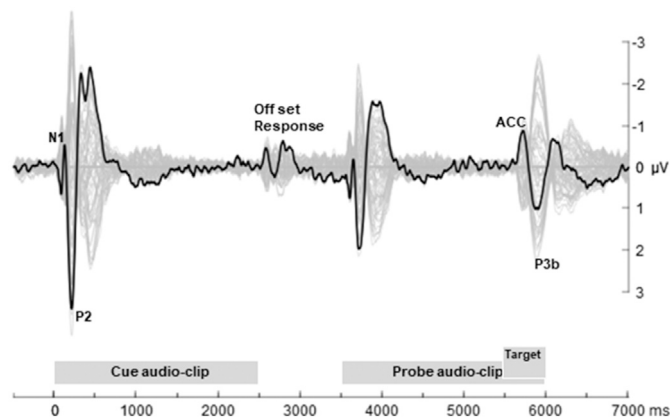

Fig. 3. Butterfly plot of group mean event-related potentials (ERPs) for the whole trial during the neutral condition. The epoch comprises 500 ms prior to and 7000 ms after the onset of cue audio clip. Grey boxes depict the sub-events. The grey lines show ERPs from all scalp electrodes. The black line shows ERPs at the midline fronto-central electrode (FCz). For illustration purposes, the data were filtered (high-pass filter = 1 Hz and low-pass filter = 20 Hz). The sustained potential was filtered out to emphasize the transient onset and offset responses of each sub-event. ACC = acoustic change complex.
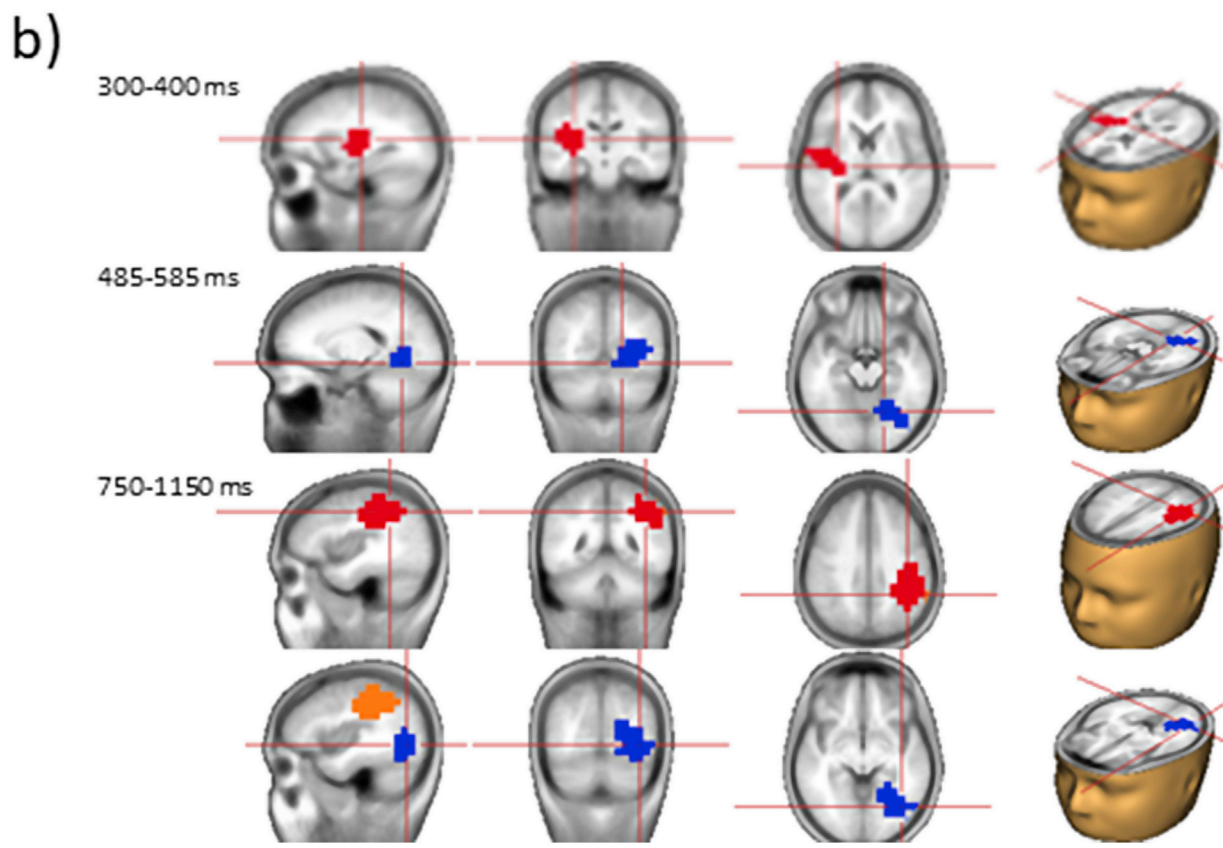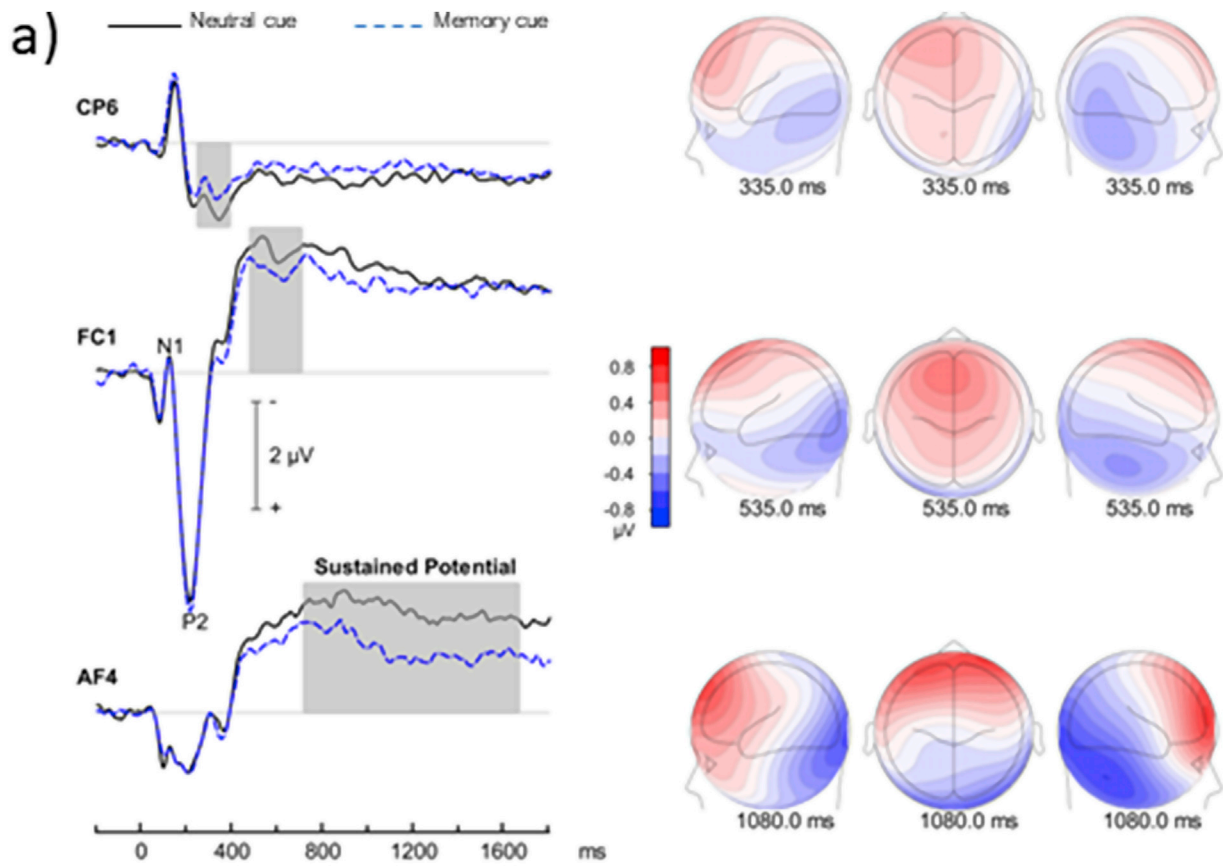
(caption on next page)

Fig. 4. a) Event-related potentials (ERPs) elicited by the cue audio clip from a representative electrode for cluster #1 (AF4, bottom row), #4 (FC1, middle row), and #6 (CP6, top row), which cover three distinct time windows. The grey shaded box highlights the interval where ERP amplitude was significantly different between the memory and neutral cues. Top right: Contour maps (left, top, and right views) show the amplitude distribution for the differences of memory minus neutral cue condition. Warm and blue colors indicate positive and negative amplitude, respectively. AF4 = right anterior frontal electrode; FC1 = left fronto-central electrode; CP6 = right central-parietal electrode. b) Difference in source activity when the target was preceded by a memory or a neutral cue. Red = greater source activity for the memory condition; blue = greater source activity for the neutral condition. For the 305–365 ms interval, the peak was located in the right temporal lobe near primary auditory cortices (Talairach coordinate: x = 31.5; y = -16.9; and z = 9.7). For the 505–565 ms interval, the difference in source activity was located in right medial-posterior temporal cortex (Talairach coordinate: x = 31.5; y = -16.9; and z = 9.7). For the 1050–1110 ms interval, there was two significant clusters. The peak of the first cluster was located in the left parietal areas (Talairach coordinate: x = -38.5; y = -44.9; and z = 37.7) whereas the peak of the second cluster was located in posterior-medial temporal cortex (Talairach coordinate: x = -24.5; y = 44.9; and z = -18.3).

Table 1

Cue audio clip: Summary of the channel level cluster-based permutation statistics.

Cluster	Latency (ms)	P value	Electrodes
1	723–1678	P < 0.001	AF3; F1; F3; F5; FT7; FC5; FC3; FC1; C3; FP2; AF8; AF4; AFz; Fz; F2; F4; F6; FC6; FC4; FC2; FCz; C2; C4
2	840–1246	P < 0.001	P1; P3; P7; P9; PO7; PO3; O1; Iz; Oz; POz; T8; TP8; CP6; P4; P6; P8; P10; PO8; PO4; O2; CB1; CB2; TP9; TP10
3	1685–1998	P = 0.007	FP2; AF8; AF4; Fz; F2; F4; LO2
4	484–715	P = 0.007	AF3; F1; F3; F5; FC5; FC3; FC1; C3; CP3; CP1; AF4; Fz; F2; FC4; FC2; FCz; C2
5	605–836	P = 0.036	C6; TP8; CP6; P8; P10; CB2; TP10
6	252–398	P = 0.041	FT8; C6; T8; TP8; CP6; P8; P10; CB2; TP10

Table 2

Cue audio clip: Summary of the channel level cluster-based permutation statistics for time-frequency analysis.

Frequency-band	Cluster Latency	P value	Electrodes
Alpha (8–12 Hz)	150–1100	P = 0.001	Fp1; AF3; F1; FC1; C1; C3; C5; TP7; CP5; CP3; CP1; P1; P5; P7; P9; PO7; PO3; O1; Pz; CPz; Fpz; Fp2; AF4; AFz; Fz; F2; F4; FC4; FC2; FCz; Cz; C2; C4; CP6; CP4; CP2; P2; P4; CB1; TP9
Beta (13–29 Hz)	400–900	P = 0.012	FC1; TP7; CP1; P1; P3; P5; P7; P9; PO7; PO3; O1; Iz; Oz; POz; Pz; CPz; Fpz; Fz; FCz; Cz; CB1; TP9; FT10; F10;
Gamma (30–50 Hz)	1550–1950	P = 0.005	AF7; FC5; FC6; C4; P2; TP9; FT9; FT10

difference coincided in time with the decay of the P3b wave and was located over the parietal scalp area contralateral to the target location (Fig. 8, Table 3). The sustained positivity was larger for the neutral than the memory cue. This difference in ERP amplitude preceded the mean RT by about 150 ms. The second cluster was located over the frontal areas ipsilateral to the target and was characterized by enhanced negativity between 260 and 326 ms after target onset for the neutral than memory cue trials.

The contrast of source activity between when the target was preceded by a memory or neutral cue revealed significant differences, with greater source activity in medial anterior temporal lobe contralateral to the target location for the memory than for the neutral cue condition during the 260–320 ms interval. For the neutral condition, the strength of source activity was inversely related to RTs such that increases source activity trended toward faster RTs (r = −0.301, ns). For the memory condition, the correlation between source strength and RT was not significant (r = −0.095). For the 430–490 ms interval, the contrast between memory and neutral cue conditions showed greater source activity in the inferior frontal gyrus ipsilateral to the target location when the target was

preceded by neutral than memory cue. The correlations between source strength and RTs were not significant for the memory nor the neutral conditions. Lastly, we compared the "memory effect" of different oscillatory activity between memory and neutral trials for left- and right-lateralized targets. Spectral power changes, however, were not significantly different for left and right targets, suggesting that the "memory effect" was little affected by target location. The lack of difference in spectral power could be due to the relatively low number of lateralized targets, making it difficult to detect a statistically reliable difference in spectral power.

5. Discussion

This study sought to clarify the nature and timing of the cognitive processes that underlie the effects of auditory LTM on target detection. Target detection was faster for a memory cue than for a neutral cue, suggesting that target-context associations formed during the learning phase could be retrieved to guide auditory attention. However, the cue-based gains in RT were less pronounced than those observed in Zimmermann et al. (2017), and the effect of the contextual cue on accuracy was not statistically significant. This discrepancy between the current study and that of Zimmermann et al. (2017) is likely due to the difference in the number of audio clips used. The larger number of audio clips used in the present study (104 vs 80 audio clips) made it more difficult for participants to learn the association between the audio clip and the target as evidenced by accuracy for the memory trials at the end of four learning blocks (73% vs. 79% from Experiment 3, Zimmermann et al., 2017). For the memore trials, the target detectability during the learning task was lower than aimed for. This is likely due to our calibration task, which used only four different audio clips. Moreover, the target was first presented at a volume that made it easy to hear. Therefore, participants could learn quickly what to listen for thereby providing thresholds that were likely lower than what would have been estimated using a larger number of audio clips.

In the present study, learning differed for memory and neutral trials, with participants showing more rapid learning for neutral than memory trials. This difference may be due to the fact that neutral trials were associated with a single decision (absence of the target) whereas in memory trials, not only did participants had to remember whether the target was present or absent, they also had to indicate whether it was at the left or right.

Our response time data are consistent with those of earlier studies showing that visual (Summerfield et al., 2006, 2011) or auditory contextual memory cues (Zimmermann et al., 2017, 2019) facilitate target detection. In the present study, the same trials were used during learning and test. That is, for memory-guided attention task, all audio clips had been heard before, with the only difference being the association, or lack thereof, between the audio clip and a lateralized target indicated by pressing the left, right, or down arrow key for left, right or neutral trials, respectively. Such a design, however, created an imbalance for neutral trials in stimulus-response mapping such that during the memory-guided attention task, participants had to indicate where the target was located by pressing the left or right button. Hence, we cannot rule out the possibility that such a change in stimulus-response mapping may increase response time for the neutral trials and could influence the
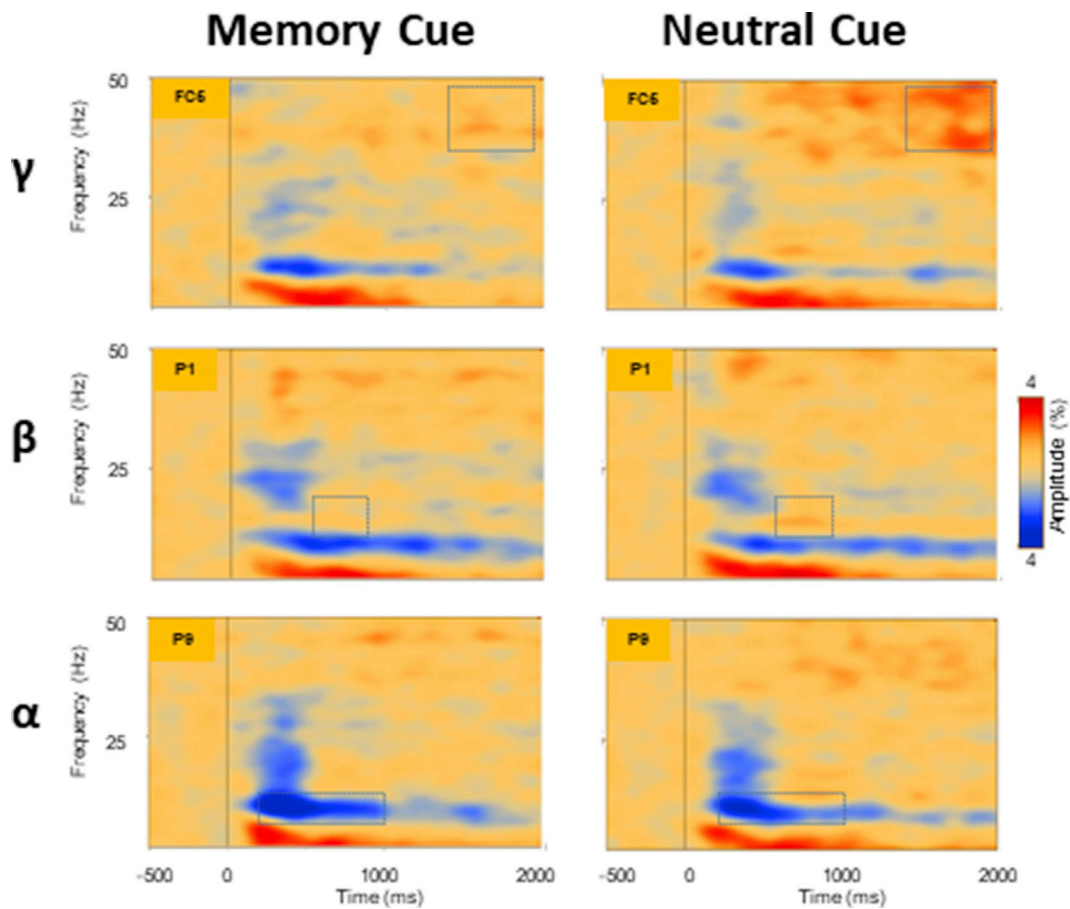
Fig. 5. Time-frequency analyses revealed desynchronization of alpha-, beta- and gamma-band activity for memory cue trials compared to neutral cue trials during the cue audio clip. The light grey box highlights the area that was significantly different between the two cueing conditions. FC5 = left fronto-central electrode; P1 = left mid parietal electrode; P9 left lateral parietal electrode.
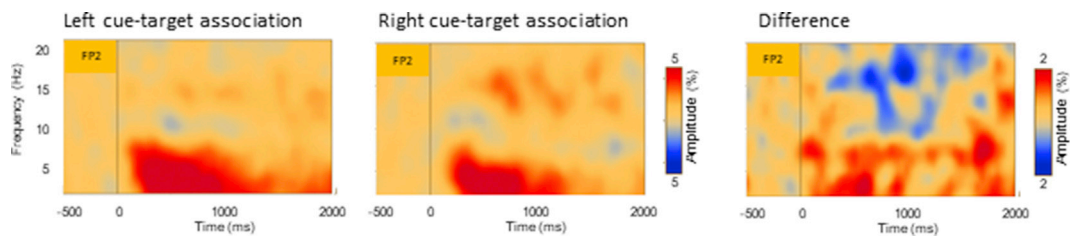

Fig. 6. Time-frequency analyses revealed desynchronization of theta as a function of the cue-target association during the cue audio clip period. FP2 = right fronto-topolar electrode.

strategy with which the participants complete the task. For instance, participants may have relied more on temporal information, which was constant between the learning phase and the test phase. Future research, controlling for stimulus-response mapping, may help address this possibility.

The mechanisms underlying memory-guided search, however, may be different between visual and auditory paradigms. Compared to orienting attention in a visual scene, memory-guided attention during auditory scene analysis may not rely as much on binding targets with a particular within-scene sound object or a memory for the spatial configuration of various sound objects. Rather, memory-guided attention may rely on the memory for the audio clip and target as a whole. Also, the behavioural benefit of auditory contextual memory may not be due to anticipatory attention to where the target might occur (spatial orienting of attention) but rather to when it might occur (temporal orienting of attention). The analyses of neuro-electric activity revealed memory-

guided attention effects that are largely consistent with the deployment and management of auditory attentional resources to when the target occurred.

5.1. ERP: cue audio-scene and memory-guided attention

The behavioural benefit of contextual auditory memory was associated with at least three distinct ERP modulations in response to the cue. The early modulation peaks at about 325 ms and coincided with enhanced activity in auditory sensory areas. The difference in ERP amplitude may index implicit perceptual memory (Wagner and Gabrieli, 1998). This interpretation is consistent with prior research showing enhanced activity in auditory areas for familiar environmental sounds (Kirmse et al., 2012), familiar voices (Birkett et al., 2007), or personal ring tones (Roye et al., 2010).

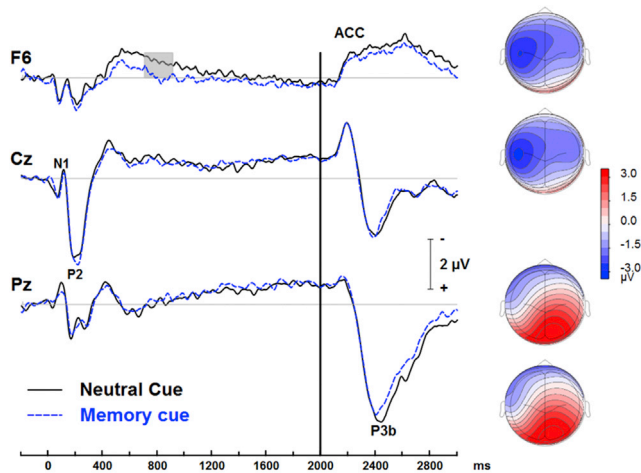Memory cues were also associated with enhanced activity in the left

Fig. 7. Event-related potentials (ERPs) elicited by the probe audio clip and target stimulus. The grey shaded box highlights the interval where ERP amplitude was significantly different between the memory and neutral cues. The vertical line indicates the target onset. The contour maps (top view) show the amplitude distribution for the acoustic change complex and the P3b wave elicited by the target sound.

Table 3
Target tone: Summary of the channel level cluster-based permutation statistics.

Cluster	Latency (ms)	P value	Electrodes
1	428–494	P = 0.033	CP1; Pz; CPz; FC2; Cz; C2; CP6; CP4; CP2; P2; P4; P6; PO4
2	260–326	P = 0.047	FP1; AF3; F1; F3; F5; FC3; FC1; AF4; AFz; Fz; F2; F6; FC2; FCz

parietal cortex. The results from our distributed source analysis are consistent with those from prior fMRI studies (Stokes et al., 2012; Summerfield et al., 2006), suggesting that the parietal cortex may play an important role in recollection and allocation of auditory attentional resources. Our findings also appear to be consistent with current models of attention, which posit that the parietal cortex plays an important role in orienting attention to representation in short-term memory (Alain et al., 2015; Backer et al., 2015; Gazzaley and Nobre, 2012) and long-term memory (Cabeza et al., 2008; Rugg and King, 2018), as well as memory-guided attention to a visual target embedded in an image of everyday scenes (Summerfield et al., 2006). Although our findings and those from the literature suggest a supramodal role of the parietal cortex in memory-guided attention, further research is needed to determine whether the parietal activity would differ when attention is oriented toward an incoming auditory or visual stimulus.

Both memory and neutral cues generated a sustained potential over the frontal scalp region. The sustained modulation was larger for neutral than for memory cues and coincided with greater activity in the medial temporal lobe. This enhanced activity for neutral cues could reflect the maintenance of the target location in working memory: in case of the neutral cue, two rather than one alternative target locations have to be maintained. The sustained potential may also index a template matching process, comparing the incoming auditory cue with those in memory. The lower amplitude for memory-cue trials may indicate quicker recognition for memory than in the neutral cues. Neutral trials, albeit also presented during the learning task, were not associated with any target, and consequently, a search through memory may take longer and be more effortful in order to determine that no target was associated with that audio clip. This may be analogous to a visual search of a cluttered display that does not contain the target. Alternatively, the presence of a cue when it was unexpected in the neutral condition may give rise to a

prediction error which often is associated with increased medial temporal lobe activity (Henson and Gagnepain, 2010) and generation of signals akin to those associated with mismatch negativity (see below).

5.2. Memory guided-attention in audition: space vs time

In the literature on visual memory-guided attention, neural activity consistently reflects anticipatory spatial biases triggered by LTM. This is the case when basic arrays are used (Chaumon et al., 2009a, 2009b; Kasper et al., 2015), or when more complex real-world scenes drive spatial attention (Stokes et al., 2012; Summerfield et al., 2011). Such anticipatory spatial biases may occur because visual contextual memory is spatially rich, as opposed to associative memory recollection of audio clip—target pairings in the current paradigm.

In the present study, the comparison of oscillatory activity during the cue audio clip period revealed greater theta power when the clip was paired with a left rather than a right auditory target. This difference in theta power may reflect an associative memory between the audio clip and the target location. This account is consistent with prior research showing a strong relationship between theta oscillation and spatial memory (Miller et al., 2013, 2018). This location-specific effect, however, was limited to theta power, with no significant difference in the alpha, beta, or gamma bands.

Here, we showed comparable ERP modulation and alpha lateralization when the memory cue was associated with a left- or right-lateralized target. Together, these findings suggest that in the present study, memory-guided attention may operate on timing rather than on spatial location. Participants may have relied more on timing information than on spatially detailed context to guide attention because the former was always constant. Prior research on the auditory attentional blink has shown that attention can be oriented to a designated time (Shen and Alain, 2011, 2012; Shen et al., 2016). Therefore, it is reasonable to assume that participants may have used the contextual cue to orient attention in time rather than in space. Also, because the auditory cortex is not as spatially organized as the visual cortex, neural activity associated with spatial attention is more difficult to detect in the auditory modality (Gamble and Luck, 2011).

Memory cues were associated with alpha and beta desynchronization, which may reflect an increase in attention required to search through memory and suppression of task-irrelevant stimuli (Foxe and Snyder, 2011; Hong et al., 2008). Our results were striking given that changes in alpha activity were analyzed over a long time interval and alpha modulation was observed over 3 s before the expected onset of the target, which was much earlier than modulations observed in visual memory-guided attention paradigms (Stokes, 2011; Summerfield et al., 2011). In these visual paradigms, alpha modulation was analyzed over brief 100 ms intervals that offset 100 ms prior to the onset of the target, and therefore necessarily reflected attention-related mechanisms in preparation for the target. Participants in our study were aware that two repeated audio clips (cue and probe audio clips) would be presented, and that the target tone was embedded towards the end of the probe audio clip. Nonetheless, we observed a desynchronization of alpha activity at central sites during the cue audio clip several seconds before the expected target onset, which could index the deployment of attention (Backer et al., 2015). The alpha and beta suppression may indicate that participants are searching through memory (Backer et al., 2015) and long-term memory retrieval (Tamura et al., 2016). This interpretation is consistent with evidence from MEG studies showing that alpha rhythm was suppressed when participants search through memory for tones in a Sternberg task (Kaufman et al., 1992; Rojas et al., 2000). The sustained alpha power reduction could also reflect the suppression of task-irrelevant information in order to "protect" information retrieved from long-term memory and into working memory (Ahveninen et al., 2017; Payne and Sekuler, 2014).

Here, it is important to note the spatial layout of the reported desynchronization of alpha and beta-band activity differs from that
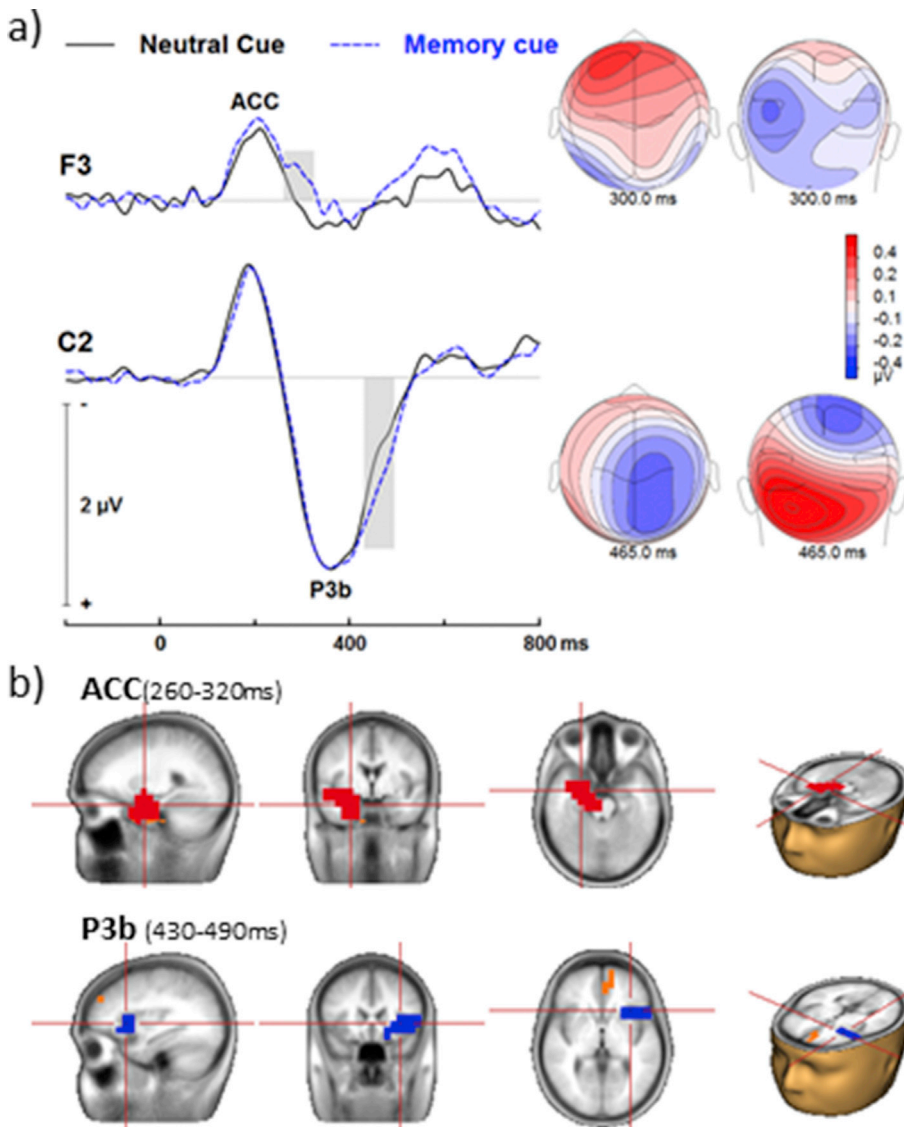
Fig. 8. a) Event-related potentials (ERPs) time-locked on target onset from a right central electrode. The ERPs were baseline corrected using a 200-ms interval before the target onset. The grey shaded box highlights the time where the memory and neutral conditions were different. The contour maps (top view) show the amplitude distribution for the difference wave. b) Difference in source activity when the target was preceded by a memory or a neutral cue. Red = greater source activity for the memory condition; blue = greater source activity for the neutral condition.

observed in visual memory-guided attention tasks, in which informative cues immediately preceding targets elicit contralateral alpha/low beta desynchronization at posterior scalp regions (i.e., most frequently PO7/8). In our paradigm, memory cues only provided information about the spatial location of the target stimulus. Although expectations for timing were generated in the learning phase and biased responses (Zimmermann et al., 2017), participants were cued spatially and asked to respond to spatial location. In addition to spatial cues, providing temporal cues to directly guide attention would allow us to assess the effects of LTM for timing on attention, and determine whether temporal and spatial cueing in audition activates separate brain structures and processes. Based on what is known about implicit expectations formed by regularities in space and time, both in vision (Doherty et al., 2005) and audition (Rimmele et al., 2011), we might expect that temporal cueing will have effects at earlier processing stages than spatial cueing. This may be the case, particularly in audition, where temporal processing is especially important.

It is important to note that the memory cues in this paradigm were lengthy, totalling over 5 s for the cue and probe audio clip presentation before target onset. Due to the dynamic nature of auditory stimuli, a longer cue was needed to allow for interpretation of the sound and recollection of the associated target, compared to 100 ms cues used in

some visual paradigms (Patai et al., 2012). A long cueing period introduced the possibility for a large variation in the timing of the memory-guided process across trials and across participants. For example, audio clips strongly encoded in memory would likely lead to earlier recollection and attentional shifts compared with weakly encoded cues. Averaging ERPs across trials and participants would then reduce the effects of the cueing condition. Therefore, an alternative explanation for the lack of lateralization in ERPs at the cueing period is the variability in the timing of the memory-guided attention process. In future studies, modifying the paradigm to reduce the necessity for long auditory cues may alleviate this problem. Alternatively, fMRI could be used to capture the spatial distribution of processing in auditory memory-guided attention.

5.3. ERP: probe and target detection

We also examined the effects of memory cues on the processing of the target itself. Targets preceded by neutral cues elicited an increase in negativity, reminiscent of a mismatch negativity (MMN) elicited by violations of prediction in an auditory scene (Picton et al., 2000). In the neutral cue condition, the incoming target may be more likely to violate some expectation whereas the target in the memory cue trial may match

the knowledge of the association between the audio clip and the target. Target stimuli also generated a P3b wave at midline parietal sites. Smaller P3b is typically interpreted as indicating the engagement of fewer attentional resources (Fritz et al., 2007; Picton, 1992). In this case, fewer attentional resources were required to process targets preceded by a memory cue, in which participants already knew (explicitly and/or implicitly) when and where the target should occur, compared to targets preceded by a neutral cue, in which participants may have had to divide attention between the left and right auditory field. Comparatively, in visual contextual cueing, targets embedded within informative contexts are marked by earlier electrophysiological changes, such as N2pc (Kuo et al., 2009; Patai et al., 2012; Summerfield et al., 2011), thought to reflect underlying changes in attention allocation. These differences in neural modulations may also be a function of task differences, in which classic contextual cueing involves quicker and more automatic orienting based on context, or it may be a result of modality differences (Fritz et al., 2007).

5.4. Memory-guided attention persists over time

Auditory LTM speeded response times to the lateralized target tone 24 h after learning. Chun and Jiang (2003) conducted a study assessing memory-guided attention over different retention delays examining implicit contextual cueing using visual arrays. They reported that participants' performance to detect targets within repeated arrays one week after learning did not differ from performance with only one day of delay. However, since the contextual cueing paradigm confounded learning and memory-guided attention, their study did not allow comparison of memory-guided attention at different delay intervals. In the current study, the learning phase was separated from the memory-guided attention task. In the memory-guided attention task, all audio clips were equally familiar to participants, but only a subset was associated with an auditory target. Our findings suggest that auditory associations can last up to 24 h after learning and can facilitate signal detection.

It is important to note the potential significance of understanding memory-guided attention to neurodegenerative disease. Though basic auditory perception is rarely impaired in Alzheimer's disease (AD), higher-order auditory impairments appear to be a hallmark of early AD (Golden et al., 2015a, 2015b; Goll et al., 2011; Miller et al., 2010). In a recent study in our laboratory using the current paradigm, we showed that auditory memory-guided attention was significantly impaired in middle-aged asymptomatic carriers of Apolipoprotein ε4, a risk gene for AD (Zimmermann et al., 2019). Impairments in contextual learning in at-risk individuals have also been reported in visual work (Negash et al., 2007a, 2007b, 2015).

Though attention and memory are predominantly tested in isolation in clinical practice, interest was recently shifting towards examining the interface of cognitive domains. Understanding memory-guided attention in audition in healthy young adults will provide the groundwork for this changing landscape.

6. Concluding remarks

In conclusion, auditory LTM facilitates signal detection. Incoming sound triggers a long-term representation, which then can be used to manage attentional resources. Behavioural gain observed in prior studies were thought to reflect anticipatory spatial attention. The analyses of neuroelectric brain activity, however, show that auditory LTM may speed response time by guiding attention to a specific point in time. This effect is long-lasting, being present at least 24 h after learning. Memory-guided attention also depends on a widely distributed neural network that comprises sensory cortices, parietal cortex and medial temporal lobe. Further research using functional magnetic resonance imaging is needed; however, to delineate more precisely the neural substrate of memory-guided attention.

Declaration of competing interest

The authors declare no competing financial or non-financial interests.

CRediT authorship contribution statement

Jacqueline Zimmermann: Conceptualization, Methodology, Project administration, Data curation, Formal analysis, Writing - original draft, Writing - review & editing. **Bernhard Ross:** Conceptualization, Data curation, Formal analysis, Writing - review & editing. **Morris Moscovitch:** Conceptualization, Data curation, Supervision, Writing - review & editing. **Claude Alain:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing - review & editing.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuroimage.2020.116979.

References

Ahveninen, J., Seidman, L.J., Chang, W.T., Hamalainen, M., Huang, S., 2017. Suppression of irrelevant sounds during auditory working memory. Neuroimage 161, 1–8.

Alain, C., Arnott, S.R., Gillingham, S., Leung, A., Wong, J., 2015. The interplay between auditory attention and working memory. In: Jolicoeur, P., Lefebre, C., Martinez-Trujillo, J. (Eds.), Mechanisms of Sensory Working Memory. Elsevier, pp. 215–228.

Alain, C., Arsenault, J.S., Garami, L., Bidelman, G.M., Snyder, J.S., 2017. Neural correlates of speech segregation based on formant frequencies of adjacent vowels. Sci. Rep. 7, 40790.

Alain, C., Cusimano, M., Garami, L., Backer, K.C., Habelt, B., Chan, V., Hasher, L., 2018. Age-related differences in orienting attention to sound object representations. Neurobiol. Aging 66, 1–11.

Backer, K.C., Binns, M.A., Alain, C., 2015. Neural dynamics underlying attentional orienting to auditory representations in short-term memory. J. Neurosci. : Off. J. Soc. Neurosci. 35, 1307–1318.

Beniczky, S., Duez, L., Scherg, M., Hansen, P.O., Tankisi, H., Sidenius, P., Sabers, A., Pinborg, L.H., Uldall, P., Fuglsang-Frederiksen, A., 2016. Visualizing spikes in source-space: rapid and efficient evaluation of magnetoencephalography. Clin. Neurophysiol. 127, 1067–1072.

Bey, C., McAdams, S., 2002. Schema-based processing in auditory scene analysis. Percept. Psychophys. 64, 844–854.

Birkett, P.B., Hunter, M.D., Parks, R.W., Farrow, T.F., Lowe, H., Wilkinson, I.D., Woodruff, P.W., 2007. Voice familiarity engages auditory cortex. Neuroreport 18, 1375–1378.

Buzsaki, G., 1989. Two-stage model of memory trace formation: a role for "noisy" brain states. Neuroscience 31, 551–570.

Buzsaki, G., 1998. Memory consolidation during sleep: a neurophysiological perspective. J. Sleep Res. 7 (Suppl. 1), 17–23.

Cabeza, R., Ciaramelli, E., Olson, I.R., Moscovitch, M., 2008. The parietal cortex and episodic memory: an attentional account. Nat. Rev. Neurosci. 9, 613–625.

Chaumon, M., Hasboun, D., Baulac, M., Adam, C., Tallon-Baudry, C., 2009a. Unconscious contextual memory affects early responses in the anterior temporal lobe. Brain Res. 1285, 77–87.

Chaumon, M., Schwartz, D., Tallon-Baudry, C., 2009b. Unconscious learning versus visual perception: dissociable roles for gamma oscillations revealed in MEG. J. Cognit. Neurosci. 21, 2287–2299.

Cherry, E.C., 1953. Some experiments on the recognition of speech with one and with two ears. J.Acousitcal.Soc. Am. 25, 975–979.

Chun, M.M., Jiang, Y., 1998. Contextual cueing: implicit learning and memory of visual context guides spatial attention. Cognit. Psychol. 36, 28–71.

Chun, M.M., Jiang, Y., 2003. Implicit, long-term spatial contextual memory. J. Exp. Psychol. Learn. Mem. Cognit. 29, 224–234.

Cohen, M.A., Horowitz, T.S., Wolfe, J.M., 2009. Auditory recognition memory is inferior to visual recognition memory. Proc. Natl. Acad. Sci. U.S.A. 106, 6008–6010.

Crutcher, R.J., Beer, J.M., 2011. An auditory analog of the picture superiority effect. Mem. Cognit. 39, 63–74.

Del Giudice, R., Blume, C., Wislowska, M., Wielek, T., Heib, D.P., Schabus, M., 2016. The voice of anger: oscillatory EEG responses to emotional prosody. PloS One 11, e0159429.

del Giudice, R., Lechinger, J., Wislowska, M., Heib, D.P., Hoedlmoser, K., Schabus, M., 2014. Oscillatory brain responses to own names uttered by unfamiliar and familiar voices. Brain Res. 1591, 63–73.

Dimitrijevic, A., Pratt, H., Starr, A., 2013. Auditory cortical activity in normal hearing subjects to consonant vowels presented in quiet and in noise. Clin. Neurophysiol. 124, 1204–1215.

Doherty, J.R., Rao, A., Mesulam, M.M., Nobre, A.C., 2005. Synergistic effect of combined temporal and spatial expectations on visual attention. J. Neurosci. : Off. J. Soc. Neurosci. 25, 8259–8266.

Dowling, W.J., Lung, K.M., Herrbold, S., 1987. Aiming attention in pitch and time in the perception of interleaved melodies. Percept. Psychophys. 41, 642–656.

Drosopoulos, S., Wagner, U., Born, J., 2005. Sleep enhances explicit recollection in recognition memory. Learn. Mem. 12, 44–51.

Foxe, J.J., Snyder, A.C., 2011. The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention. Front. Psychol. 2, 154.

Fritz, J.B., Elhilali, M., David, S.V., Shamma, S.A., 2007. Auditory attention–focusing the searchlight on sound. Curr. Opin. Neurobiol. 17, 437–455.

Gamble, M.L., Luck, S.J., 2011. N2ac: an ERP component associated with the focusing of attention within an auditory scene. Psychophysiology 48, 1057–1068.

Gamble, M.L., Woldorff, M.G., 2015. The temporal cascade of neural processes underlying target detection and attentional processing during auditory search. Cerebr. Cortex 25, 2456–2465.

Gazzaley, A., Nobre, A.C., 2012. Top-down modulation: bridging selective attention and working memory. Trends Cognit. Sci. 16, 129–135.

Golden, H.L., Agustus, J.L., Goll, J.C., Downey, L.E., Mummery, C.J., Schott, J.M., Crutch, S.J., Warren, J.D., 2015a. Functional neuroanatomy of auditory scene analysis in Alzheimer's disease. NeuroImage. Clinical 7, 699–708.

Golden, H.L., Nicholas, J.M., Yong, K.X., Downey, L.E., Schott, J.M., Mummery, C.J., Crutch, S.J., Warren, J.D., 2015b. Auditory spatial processing in Alzheimer's disease. Brain : J. Neurol. 138, 189–202.

Goll, J.C., Kim, L.G., Hailstone, J.C., Lehmann, M., Buckley, A., Crutch, S.J., Warren, J.D., 2011. Auditory object cognition in dementia. Neuropsychologia 49, 2755–2765.

Gurtubay, I.G., Alegre, M., Labarga, A., Malanda, A., Artieda, J., 2004. Gamma band responses to target and non-target auditory stimuli in humans. Neurosci. Lett. 367, 6–9.

Henson, R.N., Gagnepain, P., 2010. Predictive, interactive multiple memory systems. Hippocampus 20, 1315–1326.

Hochberg, Y., Benjamini, Y., 1990. More powerful procedures for multiple significance testing. Stat. Med. 9, 811–818.

Hong, L.E., Buchanan, R.W., Thaker, G.K., Shepard, P.D., Summerfelt, A., 2008. Beta (~16 Hz) frequency neural oscillations mediate auditory sensory gating in humans. Psychophysiology 45, 197–204.

Ishii, R., Canuet, L., Herdman, A., Gunji, A., Iwase, M., Takahashi, H., Nakahachi, T., Hirata, M., Robinson, S.E., Pantev, C., Takeda, M., 2009. Cortical oscillatory power changes during auditory oddball task revealed by spatially filtered magnetoencephalography. Clin. Neurophysiol. 120, 497–504.

Johnsrude, I.S., Mackey, A., Hakyemez, H., Alexander, E., Trang, H.P., Carlyon, R.P., 2013. Swinging at a cocktail party: voice familiarity aids speech perception in the presence of a competing voice. Psychol. Sci. 24, 1995–2004.

Kaiser, J., Heidegger, T., Lutzenberger, W., 2008. Behavioral relevance of gamma-band activity for short-term memory-based auditory decision-making. Eur. J. Neurosci. 27, 3322–3328.

Kasper, R.W., Grafton, S.T., Eckstein, M.P., Giesbrecht, B., 2015. Multimodal neuroimaging evidence linking memory and attention systems during visual search cued by context. Ann. N. Y. Acad. Sci. 1339, 176–189.

Kaufman, L., Curtis, S., Wang, J.Z., Williamson, S.J., 1992. Changes in cortical activity when subjects scan memory for tones. Electroencephalogr. Clin. Neurophysiol. 82, 266–284.

Kirmse, U., Schröger, E., Jacobsen, T., 2012. Familiarity of environmental sounds is used to establish auditory rules. Neuroreport 23, 320–324.

Klatt, L.I., Getzmann, S., Wascher, E., Schneider, D., 2018. The contribution of selective spatial attention to sound detection and sound localization: evidence from event-related potentials and lateralized alpha oscillations. Biol. Psychol. 138, 133–145.

Klatt, L.I., Schneider, D., Schubert, A.L., Hanenberg, C., Lewald, J., Wascher, E., Getzmann, S., 2020. Unraveling the relation between EEG correlates of attentional orienting and sound localization performance: a diffusion model approach. J. Cognit. Neurosci. 1–18.

Kuo, B.C., Rao, A., Lepsien, J., Nobre, A.C., 2009. Searching for targets within the spatial layout of visual short-term memory. J. Neurosci. : Off. J. Soc. Neurosci. 29, 8032–8038.

Levitt, H., 1971. Transformed up-down methods in psychoacoustics. J. Acoust. Soc. Am. 49 (Suppl. 2), 467+.

Lewald, J., Hanenberg, C., Getzmann, S., 2016. Brain correlates of the orientation of auditory spatial attention onto speaker location in a "cocktail-party" situation. Psychophysiology 53, 1484–1495.

Lim, S.J., Wostmann, M., Obleser, J., 2015. Selective attention to auditory memory neurally enhances perceptual precision. J. Neurosci. : Off. J. Soc. Neurosci. 35, 16094–16104.

Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. J. Neurosci. Methods 164, 177–190.

McClelland, J.L., McNaughton, B.L., O'Reilly, R.C., 1995. Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. Psychol. Rev. 102, 419–457.

Miller, J., Watrous, A.J., Tsitsiklis, M., Lee, S.A., Sheth, S.A., Schevon, C.A., Smith, E.H., Sperling, M.R., Sharan, A., Asadi-Pooya, A.A., Worrell, G.A., Meisenhelter, S., Inman, C.S., Davis, K.A., Lega, B., Wanda, P.A., Das, S.R., Stein, J.M., Gorniak, R., Jacobs, J., 2018. Lateralized hippocampal oscillations underlie distinct aspects of human spatial memory and navigation. Nat. Commun. 9, 2423.

Miller, J.F., Neufang, M., Solway, A., Brandt, A., Trippel, M., Mader, I., Hefft, S., Merkow, M., Polyn, S.M., Jacobs, J., Kahana, M.J., Schulze-Bonhage, A., 2013. Neural activity in human hippocampal formation reveals the spatial context of retrieved memories. Science 342, 1111–1114.

Miller, K.M., Finney, G.R., Meador, K.J., Loring, D.W., 2010. Auditory responsive naming versus visual confrontation naming in dementia. Clin. Neuropsychol. 24, 103–118.

Müller, N., Weisz, N., 2012. Lateralized auditory cortical alpha band activity and interregional connectivity pattern reflect anticipation of target sounds. Cerebr. Cortex 22, 1604–1613.

Negash, S., Boeve, B.F., Geda, Y.E., Smith, G.E., Knopman, D.S., Ivnik, R.J., Howard, D.V., Howard Jr., J.H., Petersen, R.C., 2007a. Implicit learning of sequential regularities and spatial contexts in corticobasal syndrome. Neurocase 13, 133–143.

Negash, S., Kliot, D., Howard, D.V., Howard Jr., J.H., Das, S.R., Yushkevich, P.A., Pluta, J.B., Arnold, S.E., Wolk, D.A., 2015. Relationship of contextual cueing and hippocampal volume in amnestic mild cognitive impairment patients and cognitively normal older adults. J. Int. Neuropsychol. Soc. : JINS 21, 285–296.

Negash, S., Petersen, L.E., Geda, Y.E., Knopman, D.S., Boeve, B.F., Smith, G.E., Ivnik, R.J., Howard, D.V., Howard Jr., J.H., Petersen, R.C., 2007b. Effects of ApoE genotype and mild cognitive impairment on implicit learning. Neurobiol. Aging 28, 885–893.

Orban, P., Rauchs, G., Balteau, E., Degueldre, C., Luxen, A., Maquet, P., Peigneux, P., 2006. Sleep after spatial learning promotes covert reorganization of brain activity. Proc. Natl. Acad. Sci. U.S.A. 103, 7124–7129.

Patai, E.Z., Doallo, S., Nobre, A.C., 2012. Long-term memories bias sensitivity and target selection in complex scenes. J. Cognit. Neurosci. 24, 2281–2291.

Payne, L., Sekuler, R., 2014. The importance of ignoring: alpha oscillations protect selectivity. Curr. Dir. Psychol. Sci. 23, 171–177.

Picton, T.W., 1992. The P300 wave of the human event-related potential. J. Clin. Neurophysiol. 9, 456–479 official publication of the American Electroencephalographic Society.

Picton, T.W., Alain, C., Otten, L., Ritter, W., Achim, A., 2000. Mismatch negativity: different water in the same river. Audiol. Neuro. Otol. 5, 111–139.

Picton, T.W., Alain, C., Woods, D.L., John, M.S., Scherg, M., Valdes-Sosa, P., Bosch-Bayard, J., Trujillo, N.J., 1999. Intracerebral sources of human auditory-evoked potentials. Audiol. Neuro. Otol. 4, 64–79.

Richer, F., Alain, C., Achim, A., Bouvier, G., Saint-Hilaire, J.M., 1989. Intracerebral amplitude distributions of the auditory evoked potential. Electroencephalogr. Clin. Neurophysiol. 74, 202–208.

Rimmele, J., Jolsvai, H., Sussman, E., 2011. Auditory target detection is affected by implicit temporal and spatial expectations. J. Cognit. Neurosci. 23, 1136–1147.

Rojas, D.C., Teale, P.D., Sheeder, J.L., Reite, M.L., 2000. Neuromagnetic alpha suppression during an auditory sternberg task. Evidence for a serial, self-terminating search of short-term memory. Brain research. Cognit. Brain Res. 10, 85–89.

Roye, A., Schröger, E., Jacobsen, T., Gruber, T., 2010. Is my mobile ringing? Evidence for rapid processing of a personally significant sound in humans. J. Neurosci. 30, 7310–7313.

Rugg, M.D., King, D.R., 2018. Ventral lateral parietal cortex and episodic memory retrieval. Cortex 107, 238–250.

Sekeres, M.J., Bonasia, K., St-Laurent, M., Pishdadian, S., Winocur, G., Grady, C., Moscovitch, M., 2016. Recovering and preventing loss of detailed memory: differential rates of forgetting for detail types in episodic memory. Learn. Mem. 23, 72–82.

Shen, D., Alain, C., 2011. Temporal attention facilitates short-term consolidation during a rapid serial auditory presentation task. Exp. Brain Res. 215, 285–292.

Shen, D., Alain, C., 2012. Implicit temporal expectation attenuates auditory attentional blink. PloS One 7, e36031.

Shen, D., Ross, B., Alain, C., 2016. Temporal cue modulates alpha oscillations during auditory attentional blink. Eur. J. Neurosci. 44, 1833–1845.

Shen, D., Vuvan, D.T., Alain, C., 2018. Cortical sources of the auditory attentional blink. J. Neurophysiol. 120, 812–829.

Snyder, J.S., Gregg, M.K., 2011. Memory for sound, with an ear toward hearing in complex auditory scenes. Atten. Percept. Psychophys. 73, 1993–2007.

Stokes, M.G., 2011. Top-down visual activity underlying VSTM and preparatory attention. Neuropsychologia 49, 1425–1427.

Stokes, M.G., Atherton, K., Patai, E.Z., Nobre, A.C., 2012. Long-term memory prepares neural activity for perception. Proc. Natl. Acad. Sci. U.S.A. 109, E360–E367.

Summerfield, J.J., Lepsien, J., Gitelman, D.R., Mesulam, M.M., Nobre, A.C., 2006. Orienting attention based on long-term memory experience. Neuron 49, 905–916.

Summerfield, J.J., Rao, A., Garside, N., Nobre, A.C., 2011. Biasing perception by spatial long-term memory. J. Neurosci. : Off. J. Soc. Neurosci. 31, 14952–14960.

Tamura, K., Mizuba, T., Iramina, K., 2016. Hearing subject's own name induces the late positive component of event-related potential and beta power suppression. Brain Res. 1635, 130–142.

Wagner, A.D., Gabrieli, J.D., 1998. On the relationship between recognition familiarity and perceptual fluency: evidence for distinct mnemonic processes. Acta Psychol. 98, 211–230.

Wagner, T., Axmacher, N., Lehnertz, K., Elger, C.E., Fell, J., 2010. Sleep-dependent directional coupling between human neocortex and hippocampus. Cortex; a journal devoted to the study of the nervous system and behavior 46, 256–263.

Winocur, G., Moscovitch, M., 2011. Memory transformation and systems consolidation. J. Int. Neuropsychol. Soc. : JINS 17, 766–780.

Zimmermann, J., Alain, C., Butler, C., 2019. Impaired memory-guided attention in asymptomatic APOE4 carriers. Sci. Rep. 9, 8138.

Zimmermann, J.F., Moscovitch, M., Alain, C., 2016. Attending to auditory memory. Brain Res. 1640, 208–221.

Zimmermann, J.F., Moscovitch, M., Alain, C., 2017. Long-term memory biases auditory spatial attention. J. Exp. Psychol. Learn. Mem. Cognit. 43, 1602–1615.