



Published in final edited form as:

Neuroreport. 2021 January 13; 32(2): 163–168. doi:10.1097/WNR.0000000000001570.

Auditory cortex supports verbal working memory capacity

Gavin M. Bidelman^{1,2,3}, Jane A. Brown^{1,2}, Pouya Bashivan^{4,5}

¹School of Communication Sciences & Disorders, University of Memphis, Memphis, TN, USA

²Institute for Intelligent Systems, University of Memphis, Memphis, TN, USA

³University of Tennessee Health Sciences Center, Department of Anatomy and Neurobiology, Memphis, TN, USA

⁴University of Montreal, Montreal, QC, Canada

⁵Montreal Institute for Learning Algorithms (MILA), Montreal, QC, Canada

Abstract

Working memory (WM) is a fundamental construct of human cognition. The neural basis of auditory WM is thought to reflect a distributed brain network consisting of canonical memory and central executive brain regions including frontal lobe and hippocampus. Yet, the role of auditory (sensory) cortex in supporting active memory representations remains controversial. Here, we recorded neuroelectric activity via EEG as listeners actively performed an auditory version of the Sternberg memory task. Memory load was taxed by parametrically manipulating the number of auditory tokens (letter sounds) held in memory. Source analysis of scalp potentials showed that sustained neural activity maintained in auditory cortex (AC) prior to memory retrieval closely scaled with behavioral performance. Brain-behavior correlations revealed that lateralized modulations in left (but not right) AC were predictive of individual differences in auditory WM capacity. Our findings confirm a prominent role of auditory cortex, traditionally viewed as a sensory-perceptual processor, in actively maintaining memory traces and dictating individual differences in behavioral WM limits.

Keywords

auditory cortex; EEG; embedded processing; hemispheric specialization; sustained neural activity

Introduction

Working memory (WM) is the mental process that temporarily preserves and manipulates information for later deployment to our perceptual-cognitive systems. WM operations consist of both memory retrieval and active manipulations of sensory-cognitive representations. Human WM is, however, a limited capacity system [1], capable of buffering $\sim 7 \pm 2$ items in memory store at any one time [2]. Given its ubiquity to perceptual-cognitive

Address for editorial correspondence: Gavin M. Bidelman, PhD, School of Communication Sciences & Disorders, University of Memphis, 4055 North Park Loop, Memphis, TN, 38152, TEL: (901) 678-5826, FAX: (901) 525-1282, gmbdlman@memphis.edu.

Conflict of interest: none declared

processing, defining the neural mechanisms of WM is important to understand the brain basis of this core cognitive function.

Neuroimaging studies have identified correlates of WM in canonical memory and central executive brain regions including parietal lobe, (pre)frontal areas, and hippocampus [3–5]. Indeed, cross-species studies in macaque demonstrate a critical role of mid-dorsolateral prefrontal cortex (cytoarchitectonic areas 46 and 9/46) in monitoring information in working memory (i.e., episodic process) [for review, see 6]. Moreover, different stages of WM (encoding, maintenance, and retrieval) recruit different neural circuitry [4,7]. In visual WM, the directional flow of information between sensory and frontal brain areas reverses when encoding vs. maintaining items in WM (*encoding*: occipital sensory→frontal; *maintenance*: frontal→occipital sensory), revealing feedforward and feedback modes in the same underlying brain network [7]. While lower-(sensory) and higher-level (cognitive) brain regions interact during the time course of WM processing [6,8], an outstanding question to address is the degree to which sensory cortex itself accounts for differences in WM capacity.

A persistent and dominant view is that regions beyond auditory cortex (AC) drive WM in human and non-human primates [5,9–11]. Thus, while AC is responsible for precise stimulus encoding, it may not itself participate in the active maintenance of information in a memory buffer. Retrieval and subsequent manipulation of mental objects would be orchestrated via controlled coupling between sensory cortex and frontal and/or hippocampal areas [12]. Critically, such models assume that regions beyond Heschl's gyrus maintain information, although perhaps in a different representational form [13].

On the contrary, emerging evidence suggests a substantial portion of auditory WM might indeed be supported by more automatic, lower-level processing *within* the auditory system without the need for (or minimal reliance on) higher-level brain structures [14]. Indeed, growing evidence implicates sensory cortex itself in WM processing [9]. In non-human mammals, AC neurons show sustained activity that is sensitive to WM demands [15]. Such findings suggest that auditory WM, a function traditionally viewed through a cognitive lens, might be driven not by “cognitive-memory” areas (e.g., frontal lobe), *per se*, but nascent memory representations in auditory-sensitive areas [e.g., 5].

Here, we recorded high-density EEG as listeners performed a serial auditory WM task. We varied WM demand (i.e., cognitive load) by manipulating the number of tokens (letter sounds) held in memory. Using source analysis, we examined regional activations in auditory cortex that varied with memory load and, more critically, whether neural activity prior to memory retrieval could predict listeners' subsequent behavioral performance. Our findings reveal early stages of auditory cortical processing in left hemisphere reflect a direct correlate of the auditory WM trace and robustly predict individual differences in behavioral capacity limits.

Methods

Participants

The sample included $n=15$ young adults (28 ± 3 yrs; 8 female) recruited from the University of Memphis. All had normal hearing (thresholds < 25 dB HL), reported no history of neuropsychiatric illness, and were right-handed (Edinburgh laterality index $> 95\%$). One participant's data was excluded due to excessive noisy EEG. Each gave written informed consent in compliance with a protocol approved by the University of Memphis IRB.

Stimuli and task

EEGs were recorded during a version of the Sternberg memory task that parametrically varied memory load while temporally separates *encoding*, *maintenance*, and *recall* stages of WM processing [for details, see 3]. On each trial, listeners heard a random series of alphanumeric characters from a subset of 15 English letters (Fig. 1a). Tokens were 300 ms (SOA = 1000 ms) presented at 76 dB SPL over ER-3A headphones (Etymotic Research). For each trial, we varied the length (size) of the "Memory set" 2, 4, 6, or 8 characters (random draw). A 3000 ms delay ("Retention") followed the offset of the final stimulus in which listeners retained the set in memory. Following retention, a single "Probe" character was played and participants indicated if it was among the prior memory set. Feedback was provided after 300 ms. The next trial commenced after 3.4 sec. On half the trials, the probe occurred in the set; on the other half it did not. Participants received 20 practice trials for familiarization and completed a total of 60 experimental trials per set size. The task was coded in MATLAB using the Psychophysics Toolbox [16]. We logged both accuracy (% correct) and median reaction times (RTs) for probe recall. We computed listeners' WM capacity (per set size) using the $K = S(H-F)$, where S is the number of items in the memory array, H is the hit rate, and F is the false alarm rate [1,3].

EEG recordings

EEGs were recorded from 64 channels at standard 10–10 locations. Continuous data were digitized at 500 Hz with a DC-250 Hz passband (SynAmps RT amplifiers; Compumedics Neuroscan). Electrodes on the outer canthi and the superior and inferior orbit monitored ocular movements. Impedances were < 5 k Ω . Response were common average referenced.

Preprocessing was conducted in BESA[®] Research (v7) (BESA GmbH) and FieldTrip [17]. Blinks were nullified using principal component analysis (PCA). Cleaned EEGs were filtered (0.01–20 Hz; 20th order elliptical), epoched -6000 to $+4000$ ms (where $t=0$ is probe onset), and averaged per set size and listener. This window encapsulated transient ERPs to the last memory tokens in the encoding period, sustained activity within the maintenance period of interest, and subsequent retrieval/response-related ERPs following the probe (see Fig. 2). We baselined traces between -4400 and -3700 —zeroing EEGs just prior to the final memory set token—to highlight neural activity during maintenance relative to the immediately preceding encoding period.

ERP sensor and source analysis

From channel-level waveforms, we measured ERP amplitudes within a cluster of central electrodes (C1/2, Cz, CP1/2, CPz) where load-dependent changes were prominent in initial visualizations of scalp topographies (data not shown). We then measured the mean ERP amplitude in the entire maintenance window (−3000–0 ms) within this electrode cluster (Fig. 2, *inset*).

We used Classical Low Resolution Electromagnetic Tomography Analysis Recursively Applied (CLARA) [BESA[®] (v7)] to estimate the neuronal current density underlying WM. CLARA renders focal source images by iteratively reducing the source space during repeated estimations. On each iteration (x2), a spatially smoothed LORETA image was recomputed. Voxels below 1% max amplitude were removed. Two iterations were used with a voxel size of 7 mm in Talairach space (regularization = 0.01% singular value decomposition). Group-level statistical (*t*-stat) maps were computed on full brain volumes using the ‘ft_sourcestatistics’ function in FieldTrip [17] ($P < 0.05$ mask). From individual CLARA images, we extracted source amplitudes at the midpoint of the maintenance period (i.e., latency of −1500 ms) within the spatial centroids of significant clusters within bilateral superior temporal gyrus (see Fig. 3a).

Statistical analysis

We used mixed-model ANOVAs in R (R Core team, 2018; *lme4* package) with a fixed effect of set size (subjects = random effect). We used Tukey-Kramer corrections for multiple comparisons. To assess relations between region-specific neural activity and behavior, we regressed the change in AC source activation (i.e., set size 8 – set size 2) with individuals’ maximum WM capacity for the highest load condition ($K_{load\ 8}$) [e.g., 18]. Left and right hemispheres were analyzed separately. Robust regression (bisquare weighting) was performed using the ‘fitlm’ function in MATLAB[®] 2019b (The MathWorks, Inc.).

Results

Behavioral data

Recall accuracy expectedly declined [$F_{3,39}=23.60$, $p < 0.0001$] and RTs slowed [$F_{3,39}=11.31$, $P < 0.0001$] for increasing memory load (Fig. 1b,c). Similarly, K , an unbiased measure of WM capacity, increased with set size [$F_{3,39}=48.95$, $P < 0.0001$] but with more dramatic improvements between 2 and 4 items ($t_{39}=-5.62$, $P < 0.0001$) and capacity plateauing between 6 to 8 items ($t_{39}=-2.89$, $P=0.0303$) (Fig. 1d). For 8 stimulus items, K capacity was 4.72 ± 1.07 , indicating ~4–5 items could be adequately maintained in auditory WM [19]. This is consistent with known limits to WM capacity [cf. 7 ± 2 ; 2] observed in both the visual and auditory modalities [1,3,11,18,19].

EEG data

The time course of scalp ERPs (Fig. 2) tagged distinct phases of the task including evoked peaks reflecting sensory responses to the final tokens of the stimulus array and probe. Moreover, with our baseline definition, transient activity at the end of the encoding period (i.e., ERP positivity, −4000 ms) before maintenance was similar between set sizes

[$F_{3,39}=1.64$, $P=0.19$]. Yet, strong load-dependent changes in neural activity were observed during maintenance [$F_{3,39}=4.98$, $P=0.0051$]. Scalp ERPs were larger for lower (2/4) vs. higher (6/8) memory loads [sets size 2/4 vs. 6/8: $t_{39}=3.84$, $P=0.0004$] (Fig. 2, *inset*), resembling a fatigue of neural activity in more demanding conditions.

Source ERP analysis uncovered activations distinguishing WM load in foci maxima located adjacent to AC [MNI coordinates: $AC_{left}=(-43, -19, -6)$; $AC_{right}=(51, -27, -6)$] within the lateral fissure, with coverage that included parts of superior temporal gyrus and Brodmann area BA 22 (<http://sprout022.sprout.yale.edu/mni2tal/mni2tal.html>) (Fig. 3a). These centroids were just inferior to primary auditory cortex (BA, 41) in posterior Sylvian fissure and thus, might be described as adjacent auditory parabelt/association cortex [20]. AC amplitudes varied with hemisphere [$F_{1,91}=4.76$, $P=0.0317$] and set size [$F_{3,91}=4.78$, $P=0.0038$] with no interaction [$F_{3,91}=0.65$, $P=0.58$]. Responses were stronger in left compared to right hemisphere and AC differentiated lower (2/4) vs. higher (6/8) load conditions (Fig. 3b,c).

Lastly, brain-behavioral regressions revealed listeners' maximum WM capacity (i.e., K_{load}) was predicted by load-dependent changes in source responses for left [$R^2=0.36$, $P=0.0401$] but not right AC [$R^2=0.45$, $P=0.37$] (Fig. 3d,e).

Discussion

Using EEG during active WM tasks, we show early stages of auditory cortical processing reflect a robust neural correlate of the auditory WM trace. This neural index of WM performance is dominant in left hemisphere where the degree of modulation in AC predicts individuals' behavioral WM capacity (i.e., larger ERP changes associated with larger memory store).

Source localized activity showed load-related modulations with larger sustained AC responses observed for lower- (easier) compared to higher-load (harder) set size conditions. Previous fMRI work linking sustained delay period activity to WM performance is equivocal. Some reports show enhancements [5] and others suppression [14] of cortical responses with memory load. ERP studies are similarly ambiguous finding decreases and increases [10,11,18,19] in late (> 400 ms) slow wave potentials with load level. Instead, direct recordings in animals show *both* suppression and enhancement effects in AC neurons dependent on the mnemonic context of sounds [21].

Sustained firing patterns might be not be related to WM *per se*, but rather, other concurrent mental processes for task execution including control from dorsolateral prefrontal cortex [6] (although they are unlikely contingent negative variations; see *SI Discussion*). DC potentials likely reflect higher states of arousal, alertness, and or attention [22]. Indeed, low-frequency (delta band) EEG responses are closely linked to sustained attention [23]. Interestingly, suppression of sustained responses during the delay period is reduced in individuals that rely on a covert rehearsal strategies [14]. Thus, stronger AC responses for lower set sizes may reflect more covert rehearsal by our participants to refresh short-term memory representations from being overwritten and protecting memory items from decay. This

interpretation is consistent with our brain-behavior correlations, which showed listeners with stronger reduction (cf. suppression) of sustained AC activity had higher K capacity limits (e.g., Fig. 3d). Larger neural responses for lower set sizes could reflect a stronger deployment of attention in easier, less taxing conditions.

Seminal studies in primates and fMRI in humans have demonstrated that neurons maintain a representation of the stimulus via sustained activity in prefrontal cortex which outlasts the eliciting event [8]. Sustained maintenance activation is observed across multiple brain regions in tasks involving various sensory modalities [3,13]. In agreement with prior neuroimaging WM studies [10,18,19], we similarly found sustained EEG activity during WM maintenance is modulated by stimulus load. Our results corroborate fMRI findings [5,9] by confirming successful auditory WM depends on elevated activity within areas proximal to auditory cortex. While EEG certainly cannot resolve the subdivisions of AC (e.g., core, belt, parabelt), our ERP responses localized to areas of the Sylvian (lateral) fissure. Thus, our sources might be best described as stemming from adjacent auditory parabelt/association cortex [20]. Still, this is earlier than the posterior superior temporal gyrus activations typically described in auditory WM tasks [18]. These structures are typically associated with sensory-perceptual processing rather than higher cognition [12,13]. Thus, our data support an “embedded” rather than “buffer” account of WM, whereby auditory cortex acts to both encode perceptually relevant sound features but also functions as memory buffer to maintain information over short periods [cf. 13].

We also found a stark hemispheric asymmetry in how AC predicts WM capacity. “High modulators” of left hemisphere AC activity showed superior auditory WM performance. In contrast, right hemisphere did not predict behavior. fMRI studies similarly show left AC activation positively correlates with behavioral performance in WM tasks [5]. Thus, while auditory activity is bilateral during WM [current study; 5,9,10,18], left hemisphere best correlates with behavioral capacity. Leftward laterality is perhaps expected given the well-known dominance of left cerebral hemisphere to auditory-linguistic processing [24]. Indeed, both children and adults show strong brain asymmetry in WM organization, with greater rightward bias for spatial WM and leftward bias for verbal WM [25]. Given our WM task required covert verbal labeling, the stronger activation and association we find between left AC and behavior is consistent with the left hemisphere dominance of auditory-linguistic processing [24].

Acknowledgements:

Work supported by the National Institute on Deafness and Other Communication Disorders of the National Institutes of Health (R01DC016267).

References

- [1]. Cowan N The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav Brain Sci* 2001; 24 (1):87–114; discussion 114–185. [PubMed: 11515286]
- [2]. Miller GA. The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychol Rev* 1956; 63 (2):81–97. [PubMed: 13310704]

- [3]. Bashivan P, Bidelman GM, Yeasin M. Spectrotemporal dynamics of the EEG during working memory encoding and maintenance predicts individual behavioral capacity. *Eur J Neurosci* 2014; 40 (12):3774–3784. [PubMed: 25288492]
- [4]. Karlsgodt KH, Shirinyan D, Van Erp TGM, Cohen MS, Cannon TD. Hippocampal activations during encoding and retrieval in a verbal working memory paradigm. *Neuroimage* 2005 pp. 1224–1231.
- [5]. Kumar S, Joseph S, Gander PE, Barascud N, Halpern AR, Griffiths TD. A brain system for auditory working memory. *J Neurosci* 2016; 36 (16):4492–4505. [PubMed: 27098693]
- [6]. Petrides M The mid-dorsolateral prefronto-parietal network and the eopptic process In: Stuss DT, Knight RT, editors. *Principles of Frontal Lobe Function*. New York, NY: Oxford University Press; 2013 pp. 79–89.
- [7]. Bashivan P, Bidelman GM, Yeasin M. Modulation of brain connectivity by cognitive load in the working memory network. *Proceedings on the IEEE Symposium Series on Computational Intelligence (IEEE SSCI) Orlando, FL* 2014.
- [8]. Courtney SM, Ungerleider LG, Keil K, Haxby JV. Transient and sustained activity in a distributed neural system for human working memory. *Nature* 1997; 386 (6625):608–611. [PubMed: 9121584]
- [9]. Grady CL, Yu H, Alain C. Age-related differences in brain activity underlying working memory for spatial and nonspatial auditory information. *Cereb Cortex* 2008 pp. 189–199. [PubMed: 17494060]
- [10]. Huang Y, Matysiak A, Heil P, König R, Brosch M. Persistent neural activity in auditory cortex is related to auditory working memory in humans and nonhuman primates. *eLife* 2016; 5:e15441. [PubMed: 27438411]
- [11]. Lefebvre C, Vachon F, Grimault S, Thibault J, Guimond S, Peretz I, et al. Distinct electrophysiological indices of maintenance in auditory and visual short-term memory. *Neuropsychologia* 2013; 51 (13):2939–2952. [PubMed: 23938319]
- [12]. Boran E, Fedele T, Klaver P, Hilfiker P, Stieglitz L, Grunwald T, et al. Persistent hippocampal neural firing and hippocampal-cortical coupling predict verbal working memory load. *Science advances* 2019; 5 (3):eaav3687. [PubMed: 30944858]
- [13]. Yue Q, Martin RC, Hamilton AC, Rose NS. Non-perceptual regions in the left inferior parietal lobe support phonological short-term memory: Evidence for a buffer account? *Cereb Cortex* 2019; 29 (4):1398–1413. [PubMed: 29522178]
- [14]. Linke AC, Vicente-Grabovetsky A, Cusack R. Stimulus-specific suppression preserves information in auditory short-term memory. *Proc Natl Acad Sci USA* 2011; 108 (31):12961–12966. [PubMed: 21768383]
- [15]. Sakurai Y Involvement of auditory cortical and hippocampal neurons in auditory working memory and reference memory in the rat. *J Neurosci* 1994; 14 (5 Pt 1):2606–2623. [PubMed: 8182430]
- [16]. Brainard DH. The psychophysics toolbox. *Spat Vis* 1997; 10 (433–6).
- [17]. Oostenveld R, Fries P, Maris E, Schoffelen JM. Fieldtrip: Open source software for advanced analysis of meg, eeg, and invasive electrophysiological data. *Comput Intell Neurosci* 2011; 2011:1–9. [PubMed: 21837235]
- [18]. Grimault S, Nolden S, Lefebvre C, Vachon F, Hyde K, Peretz I, et al. Brain activity is related to individual differences in the number of items stored in auditory short-term memory for pitch: Evidence from magnetoencephalography. *Neuroimage* 2014; 94:96–106. [PubMed: 24642285]
- [19]. Vogel EK, Machizawa MG. Neural activity predicts individual differences in visual working memory capacity. *Nature* 2004; 428 (6984):748–751. [PubMed: 15085132]
- [20]. Petrides M, Pandya DN. Efferent association pathways originating in the caudal prefrontal cortex in the macaque monkey. *J Comp Neurol* 2006; 498 (2):227–251. [PubMed: 16856142]
- [21]. Scott BH, Mishkin M, Yin P. Neural correlates of auditory short-term memory in rostral superior temporal cortex. *Curr Biol* 2014; 24 (23):2767–2775. [PubMed: 25456448]
- [22]. Kovac S, Speckmann EJ, Gorji A. Uncensored EEG: The role of DC potentials in neurobiology of the brain. *Prog Neurobiol* 2018; 165-167:51–65. [PubMed: 29428834]

- [23]. Kirmizi-Alsan E, Bayraktaroglu Z, Gurvit H, Keskin YH, Emre M, Demiralp T. Comparative analysis of event-related potentials during Go/NoGo and CPT: Decomposition of electrophysiological markers of response inhibition and sustained attention. *Brain Res* 2006; 1104 (1):114–128. [PubMed: 16824492]
- [24]. Hickok G, Poeppel D. The cortical organization of speech processing. *Nat Rev Neurosci* 2007; 8 (5):393–402. [PubMed: 17431404]
- [25]. Thomason ME, Race E, Burrows B, Whitfield-Gabrieli S, Glover GH, Gabrieli JDE. Development of spatial and verbal working memory capacity in the human brain. *J Cogn Neurosci* 2009; 21 (2):316–332. [PubMed: 18510448]

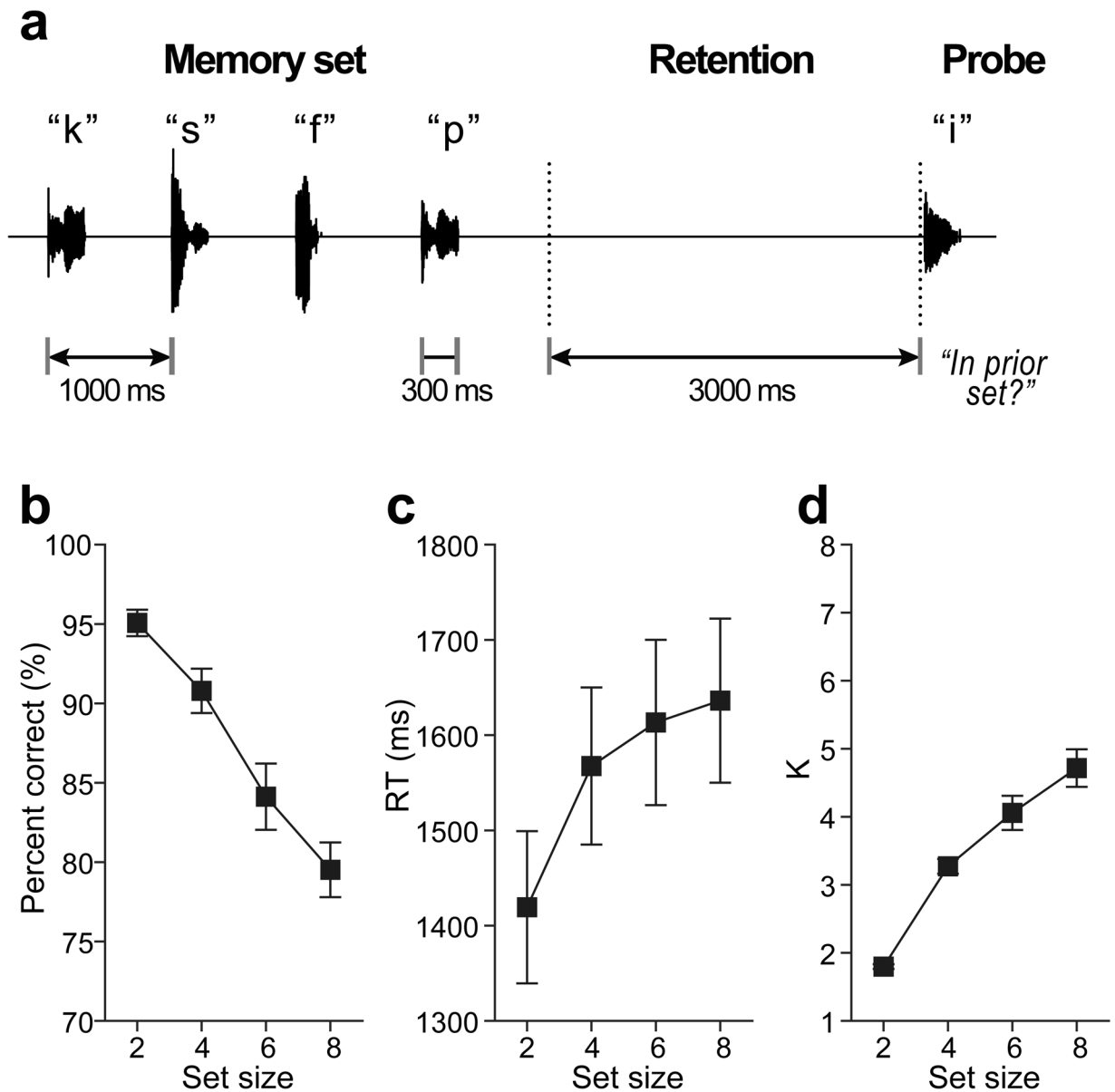


Fig. 1: Auditory WM stimulus paradigm and behavioral data.

(a) Listeners heard between 2 and 8 characters ("Memory set") presented auditorily. Following a retention period, they indicated whether a "Probe" occurred in the prior memory set. Shown here is a "no match" trial. (b-c) Behavioral accuracy for probe recall decreases and response times increase with additional memory load. (d) Working memory capacity (K) increases for small set sizes but saturates >4 items, above listeners' WM capacity limits. errorbars = ± 1 s.e.m.

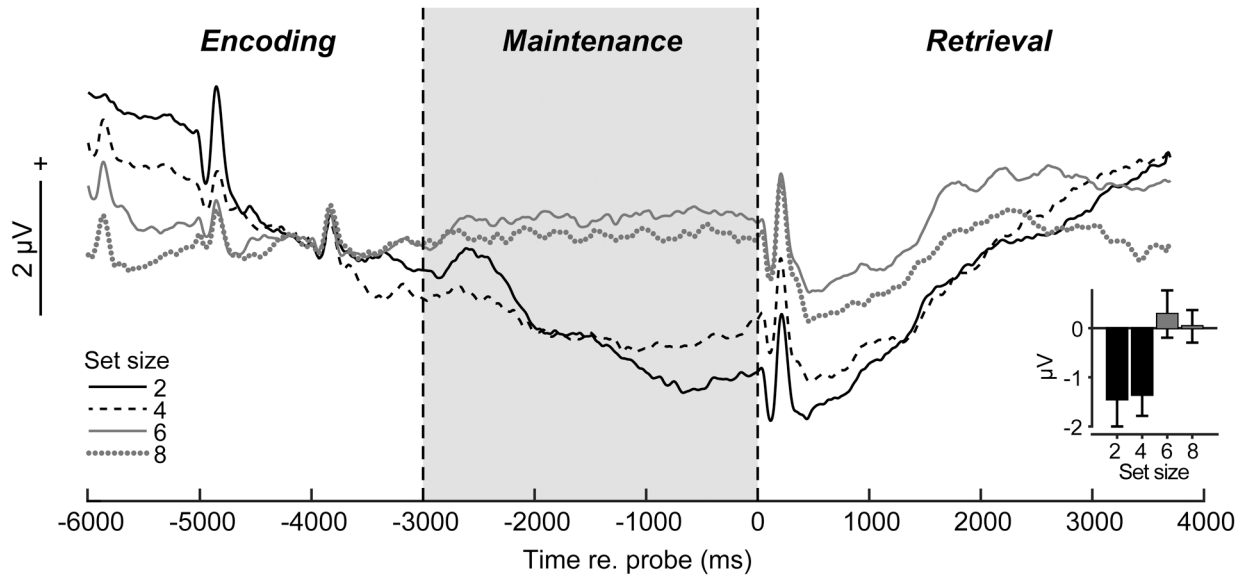


Fig. 2: Scalp ERPs reveal load-dependent modulations in sustained neural activity during WM maintenance.

(a) ERP time courses at central scalp locations (mean of electrodes C1/2, Cz, CP1/2, CPz; baseline = [-4400 to -3700 ms]). Transient peaks during “encoding” reflect auditory responses to final stimulus tokens in the memory set. Sustained activity is modulated in the 3 sec maintenance interval during memory retention (highlighted segment) and is stronger for low (2/4) vs. high (6/8) load (inset). errorbars = ± 1 s.e.m.

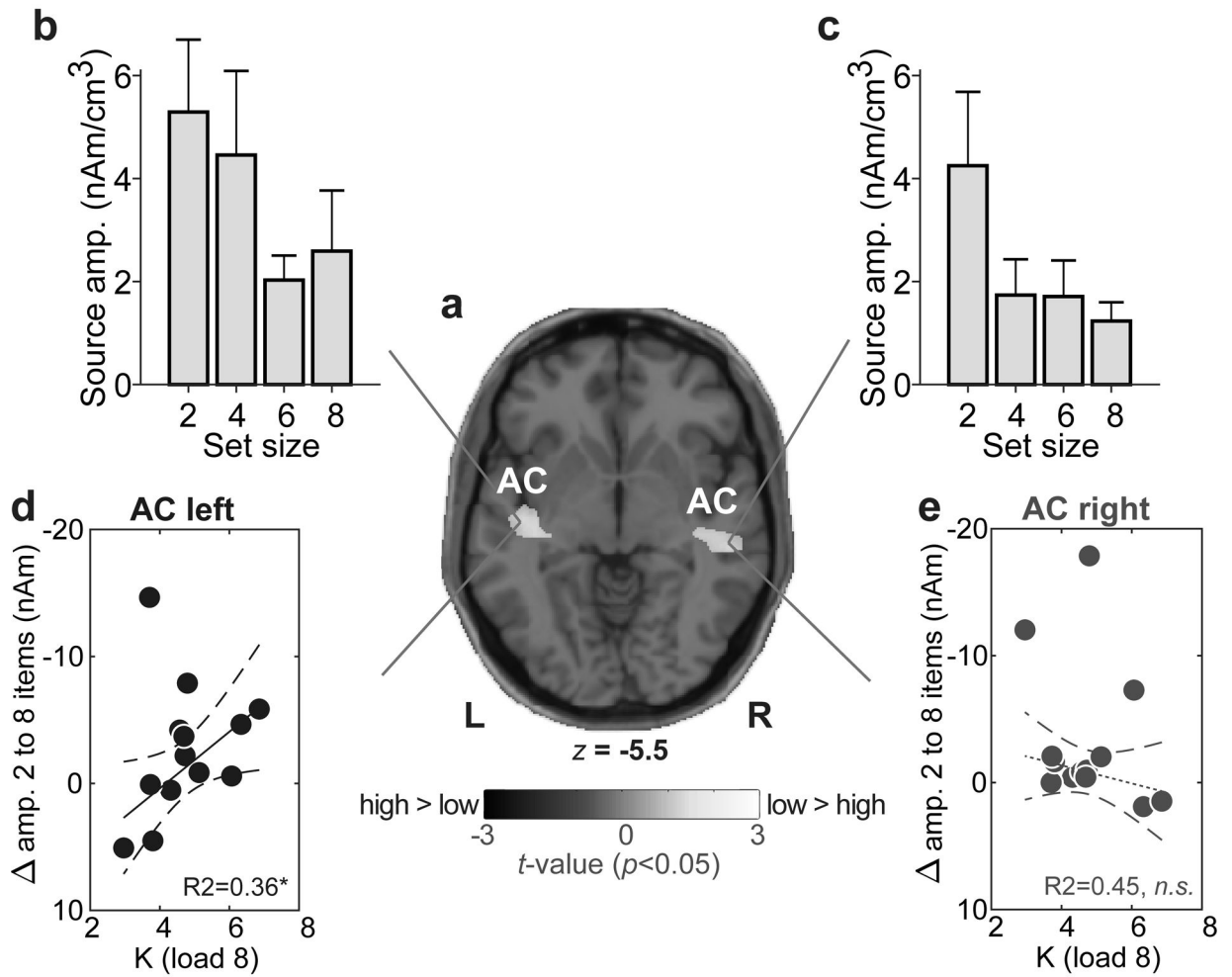


Fig. 3: Sustained neural activity maintained in AC predicts behavioral auditory WM capacity. (a) T-stat map contrasting low (2/4) vs. high (6/8) CLARA source activation maps ($P < 0.05$ masked, uncorrected). Functional data are overlaid on the MNI brain template. WM load is distinguished in bilateral Sylvian fissure. (b-c) AC amplitudes vary with set size but load-related changes in left (but not right) AC mirror the pattern observed in scalp EEG (cf. Fig. 2a) (d-e) Maximum behavioral WM capacity ($K_{load\ 8}$) is predicted by AC activity in left hemisphere; individuals with larger changes in source amplitudes with set size have larger WM capacity. No brain-behavior relation is observed in right hemisphere. Dashed lines=95% CI; solid lines, significant correlation; dotted lines, n.s. correlation. * $P < 0.05$. errorbars = ± 1 s.e.m.; AC, auditory cortex.