



Published in final edited form as:

Nat Neurosci. 2018 June ; 21(6): 808–810. doi:10.1038/s41593-018-0148-7.

Attention improves memory by suppressing spiking-neuron activity in the human anterior temporal lobe

John H. Wittig Jr.¹, Anthony I. Jang¹, John B. Cocjin¹, Sara K. Inati², and Kareem A. Zaghloul^{1,*}

¹Surgical Neurology Branch, NINDS, National Institutes of Health, Bethesda, MD 20892, USA

²Office of the Clinical Director, NINDS, National Institutes of Health, Bethesda, MD 20892, USA

Abstract

We identify a memory-specific attention mechanism in the human anterior temporal lobe, an area implicated in semantic processing and episodic memory formation. Spiking neuron activity is suppressed and becomes more reliable in preparation for verbal memory formation. Intracranial EEG signals implicate this region as a source of executive control for attentional selection. Consistent with this interpretation, its surgical removal causes a significant memory impairment for attended words relative to unattended words.

Attention mechanisms for vision are divided into modulation and selection processes¹. Modulation processes occur in early visual areas, where attention increases neural spike rate, sharpens tuning curves, and improves signal-to-noise². In contrast, selection processes occur in fronto-parietal attention areas, which exert top-down control over whichever visual area is responsible for processing the attended visual feature or location³. A key distinction is that modulation processes are found in brain areas that are visually responsive irrespective of attentional state, whereas selection processes can be found in brain areas that respond only to attentional cues⁴.

To identify attention mechanisms that enhance memory, we recorded intracranial EEG (iEEG) in 18 epileptic neurosurgery patients as they memorized words that were cued by a row of asterisks (preparatory cue; Fig. 1a,b; Supplementary Fig. 1). Cued words were remembered significantly more often than uncued words. To isolate attention signals, we contrasted high frequency power (70–200 Hz) for cued words that were successfully remembered versus uncued words that were inadvertently remembered. We evaluated the preparation phase preceding a to-be-remembered word, and the encoding phase when the

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: http://www.nature.com/authors/editorial_policies/license.html#terms

*Correspondence to: kareem.zaghloul@nih.gov.

AUTHOR CONTRIBUTIONS

J.H.W. and K.A.Z. conceptualized the study and wrote the paper. J.H.W. analyzed the data. J.H.W., A.I.J., and J.B.C. collected behavioral data and processed the single-unit data. J.B.C. localized iEEG electrodes. S.K.I. oversaw iEEG data acquisition and provided clinical assessment of iEEG waveforms and seizure focus localization. K.A.Z. performed all surgical procedures and supervised the study.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interest.

word is read, interpreted, and memorized (Fig. 1c). During the preparation phase, high frequency power in the anterior temporal lobe was significantly lower for cued words relative to uncued words (Fig. 1d, ATL; Supplementary Figs. 2–3). In contrast, during the encoding phase the anterior temporal lobe did not signal attention, though two other brain areas showed increased high frequency power during cued words (Fig. 1e; Supplementary Fig. 4): the posterior temporal lobe (PTL) and frontal cortex.

The time course of high-frequency power changes in the anterior temporal lobe revealed a robust and significant decrease following the preparatory cue in individual electrodes and across the population of participants (Fig. 1f, red traces). For uncued words, there was no change in high frequency power before, during, or after word presentation (Fig. 1f, black traces), even though the words were subsequently remembered. The lack of a response to uncued words suggests that this area is not involved in perceptual processing of written words, nor verbal memory formation, per se. Instead, the clear response to the preparatory cue implicates this region as a source of attentional selection processes that enhance memory formation. The dynamic, attention-related response was significantly attenuated preceding cued words that were subsequently forgotten, implicating a causal role for this signal in memory formation (Fig. 1i).

In the posterior temporal lobe, both cued and uncued words triggered a significant increase in high frequency power when the word was on the screen (Fig. 1g, black and red traces are greater than zero). However, power was significantly greater for cued words relative to uncued words. This response illustrates the modulatory effects of attention in visual processing and working memory^{1,5}. In contrast, in the frontal attention area, only cued words triggered an increase in high frequency power, whereas there was no change in high frequency power for uncued words (Fig. 1h). The lack of a response to uncued words, combined with a large response to cued words, is consistent with this region's role in attentional selection¹. For both regions, the response was significantly less for cued-words that were subsequently forgotten (Fig. 1j,k), consistent with their established roles in visual attention. We tested whether the magnitude of preparatory decreases in the anterior temporal lobe on any given trial predicted the subsequent magnitude of encoding increases in the posterior temporal or frontal attention areas (Supplementary Fig. 5). Indeed, the trial-by-trial relationships are significantly negatively correlated, consistent with the anterior temporal lobe being an attentional selection region that participates in a larger attention network for enhancing memory formation.

We had the unique opportunity to examine spiking-neuron responses in the anterior temporal lobe in four participants (Fig. 2a). An example neuron showed no significant change in its spike rate of 2 sp/s when the participant viewed uncued words that were subsequently remembered (Fig. 2b, black trace and raster). However, when the preparatory cue appeared, spike rate significantly decreased for 500 ms (Fig. 2b, red trace and raster). This example neuron's response mirrored that of the recorded population of 197 units. Population spike rate decreased significantly following the preparatory cue and remained low during the first 500 milliseconds of the encoding period (Fig. 2c, red versus black line).

Visual attention improves signal-to-noise in spiking neurons, measured as decreased fanofactor for individual neurons, and changes in noise correlations among pairs of neurons^{2,6}. We did not observe any systematic changes in noise correlations related to attention in the anterior temporal lobe (Supplementary Fig. 6). However, we observed a robust decrease in fanofactor immediately following the preparatory cue (Fig. 2d, red vs black trace). The preparatory decrease in fanofactor overlaps in time with the preparatory decrease in spike rate, suggesting that these processes are related. We confirmed that this was the case across the population of neurons, as we found a significant correlation between a neuron's change in spike rate and its change in fanofactor during the 500 ms preceding word onset (Fig. 2e). This suggests that the wide-spread suppression of high-frequency power across the anterior temporal lobe measured by iEEG electrodes reflects both the suppression of neural spiking activity and improved signal-to-noise.

As a causal test of anterior temporal lobe's involvement in enhancing verbal memory, we trained and tested all participants before, and 3 months after, its surgical removal (Fig. 3a,b). Removal of the anterior temporal lobe caused a significant decrease in memory performance, with cued-word recognition significantly more impaired than uncued recognition (Fig. 3c). The decrease in cued-word recognition was statistically indistinguishable for language dominant (dashed lines) and non-dominant (solid lines) temporal lobectomies (Supplementary Fig. 7). These results implicate a critical role for preparatory suppression of the anterior temporal lobe — making the semantic concepts and meanings of our experiences more salient for improved memory formation.

ONLINE METHODS

Participants

Eighteen neurosurgery patients (5 male; age 35 ± 3 years; FSIQ 99 ± 5 ; mean \pm s.e.m.; see Supplementary Table 1 and Life Sciences Reporting Summary) with drug-resistant epilepsy underwent a surgical procedure in which platinum recording electrodes were implanted beneath the dura on the cortical surface to localize the source of their seizures. In all cases, the clinical team determined the placement of the contacts to best localize epileptogenic regions. Four participants additionally and specifically consented to have micro-electrode arrays implanted in the anticipated site of surgical resection (details below). Data were collected at the Clinical Center at the National Institutes of Health (NIH; Bethesda, MD). The research protocol was approved by the Institutional Review Board (ClinicalTrials.gov Identifier NCT01273129), and informed consent was obtained from the participants. All testing took place in the participant's room during their stay in the Epilepsy Monitoring Unit or in the outpatient clinic.

We note that all behavioral and electrophysiological findings reported here come from participants with epilepsy, and thus results must be interpreted cautiously in terms of normal brain function. There are at least two ways that epilepsy could affect the interpretation of our results: first, epileptic activity itself could contaminate physiological signals, and second, the long-term malfunction of epileptic tissue could result in massive restructuring of the affected nervous system and potentially adjacent areas. Although we cannot eliminate these potential confounds, we took steps to mitigate them. First, we only analyzed electrophysiological

signals collected outside the ictal onset zone. The onset zone was determined by the clinical team using iEEG electrodes, and all micro-electrode arrays were at least 2 cm from the closest iEEG contact that detected ictal or inter-ictal activity. Second, we removed any trials showing epileptic activity (or other transient noise) in the remaining electrodes using an automated algorithm tuned to match the assessment of the clinical team. Third, the location and cellular abnormality that causes epilepsy varies across participants, thus by aggregating data across participants we effectively average out participant-specific abnormalities. Fourth, all measures reported here are a contrast between cued and uncued conditions for each participant, electrode, or neuron. Thus, any baseline differences in memory capacity (for participants), tissue abnormality (for electrodes), or noise levels (for electrodes and units) should be effectively subtracted out before statistical comparison of the effects of attention.

Behavior: Task Design

We designed a verbal memory task to isolate the neural mechanisms of attention that lead to improved memory formation (Supplementary Fig. 1). Participants were instructed to memorize words that were cued by a row of asterisks. Following a distraction task, we tested recognition memory of the cued and uncued words. Task difficulty was adjusted for each participant to achieve two performance criteria: recognition of cued words was better than that of uncued words, and recognition of uncued words was better than chance. These two criteria enable the dissociation of attention and memory mechanisms because we can directly contrast the neurophysiological signature of successful memory formation with and without attentional enhancement.

Successful memorization of cued and uncued words in our task is likely supported by episodic memory processes of familiarity and recollection⁷, though it is possible that short-term memory processes such as working memory could be involved¹. Working memory could be used to sub-vocally rehearse cued and uncued words throughout the memorization, distraction, recognition, and recall phases of the task. Although this is a possibility, we think it is unlikely that this would account for recognition memory performance in this task. First, the time from encoding to recall was between 60 to 600 seconds, depending on list length, which is longer than the intervals commonly used for working memory tests⁸. Second, during this interval participants actively engaged in several different mental processes, such as arithmetic and responding to the recognition tests, which would make it difficult to actively focus on and sub-vocally rehearse the words. Third, performance on the recall test indicates they followed our instruction to specifically memorize the cued words, which suggests the participants were not expending effort to actively rehearse uncued words that were successfully recognized.

Behavior: Performance Criteria

We adjusted task parameters on a session-by-session basis to achieve two performance criteria: (i) recognition of cued words was significantly better than uncued words, (ii) recognition of uncued words was significantly better than chance (see Supplementary Fig. 8). We defined chance as the false-alarm rate for novel foil words in seen/unseen recognition tests, or 25% for 4-alternative forced choice recognition tests. We tested for significant differences using a chi-squared test on the contingency table of seen and unseen responses⁹.

In the case of 4-alternative forced choice, we deemed uncued recognition significantly greater than chance if the 95% binomial confidence interval¹⁰ exceeded 25%.

We often dedicated the first one or more sessions to identifying the ideal task parameters while we trained the participants on the task. During training participants were instructed to say the cued words out loud during the encoding phase so we could confirm they were correctly identifying those words. After training we instructed them to remain silent except for the recall test at the end of each list. Sessions used to train the participant were discarded from all subsequent analysis. By tuning the task difficulty to each individual participant, we were able to collect iso-performance data across participants who showed a wide range of natural aptitude for the task.

Behavior: Effect of Surgery

Only participants that passed our performance criteria on a seen/unseen recognition test during pre-operative testing were included in our analysis of the behavioral effects of surgery. For each participant in this analysis, task parameters (e.g., list length) and anti-epileptic pharmaceutical treatment were identical in the pre-operative and post-operative test sessions. Pre-operative testing sessions were preceded by 1 to 3 training sessions used to identify ideal task parameters. Post-operative testing sessions were preceded by approximately 15 minutes of training to re-familiarize the participant with the task. We quantified behavioral performance in these pre- and post-operative sessions using D-prime. This measure of recognition accuracy incorporates false alarm rates to account for an individual's bias towards seen or unseen responses. We calculated D-prime as the z-transformed probability of responding "seen" to previously seen words minus the z-transformed probability of responding "seen" to a foil word, adding 0.5 the numerator and 1.0 to the denominator of the proportion of seen trials before z-scoring to protect against extreme values¹¹. The location of tissue removed from each participant was determined from post-operative MRI and was projected to a standard brain to visualize overlap of removed tissue between participants.

Intracranial EEG: Data Processing

We recorded intracranial EEG (iEEG) data from subdural electrodes (PMT Corporation, Chanhassen, MN) sampled at 1000 Hz using a Nihon Kohden EEG data acquisition system. Subdural contacts were arranged in both grid and strip configurations with an inter-contact spacing of 5 or 10 mm. Electrode localization was accomplished by co-registering the post-operative CTs with the post-operative MRIs using both FSL Brain Extraction Tool (BET) and FLIRT software packages. Pre-operative MRIs were used when post-operative MRIs were not available. The electrode locations were projected to the cortical surface of a Montreal Neurological Institute N27 standard brain, and atlas information for each electrode was obtained using a Talairach daemon^{12,13}.

Our first step in cleaning the iEEG signals was to eliminate any electrodes identified by the clinical teams as being part of the ictal onset zone during the monitoring period. Taking the remaining electrodes, we removed any additional channels and/or trials from our analysis that showed additional signs of noise (epileptic activity, physical movement of the

participant, or external sources of transient electrical perturbations) using a procedure we adapted from the EEG analysis software package FieldTrip¹⁴. For each channel and trial, we computed the variance of the voltage trace during the 5.25 seconds surrounding the encoding period (-1.5 seconds prior to word onset to 3.75 seconds following word onset). This resulted in a two-dimensional matrix of variance measures, channels by trials, from which we identified the maximum variance for each trial and the maximum variance for each channel. The procedure then calculated the quartiles of the resultant distribution of trial variances and channel variances, and identified whether any trials or channels exceeded a threshold of the third quartile + $w_{\text{thresh}} * \text{the inter-quartile range}$, where w_{thresh} is a user-specified parameter. If any trials exceeded the threshold, the maximum variance trials were iteratively removed, which the quartiles being recalculated at each step. If any of the removed trials caused a change in the number of channels that exceeded threshold, that trial was flagged for exclusion. Otherwise, if any channels exceeded threshold after all noisy trials were removed, those noisy trials were put back into the matrix, and instead noisy channels were iteratively removed. This procedure of identifying noisy trials, and then channels, was iterated until all trials and channels were within the threshold limits. When adapting this procedure to our dataset, we found that a w_{thresh} of 0.5 (equivalent to 2 standard deviations from the mean) was conservative in catching time points judges as noisy by the epileptologist on our clinical team. After systematically removing noisy channels and trials, we removed the effects of a common reference from the voltage traces by subtracting a global common average, computed across all channels separately for each trial and time point. We confirmed that any potential noise artifacts caused by saccadic eye movements do not account for the attention-related effects reported in the anterior temporal lobe (Supplementary Fig. 9).

Intracranial EEG: Power Analysis

We estimated a continuous time measure of high-frequency power using a wavelet decomposition (complex Morlet kernel; wave number 6; 70 to 200 Hz with 5 taps per octave). This frequency range is a proxy for local neural spiking activity¹⁵. We log-transformed the power estimates¹⁶, down-sampled in time using a sliding 100 ms boxcar window with a step of 50 ms, and baseline-corrected and z-transformed each time point by subtracting the mean and dividing by the standard deviation of power measured during a baseline period. Baseline was the 500 ms following the offset of the orientation cue (“+”) for all trials. We did this separately for each recording session to account for changes in day-to-day signal quality. High frequency power was computed as the average z-transformed value from all wavelets between 70 and 200 Hz, yielding a single time series for every electrode and every trial. For individual electrodes, we tested for significant differences between the cued and uncued time series by calculating a two-sample t-test across trials at every time point. For population-average electrode effects in predefined regions of interest (Supplementary Fig. 10), we first calculated a mean cued and mean uncued time series for each electrode, then averaged the time series from all electrodes in that region for each participant, and then calculated a paired t-test across participants at every time point. For individual electrodes and population averages we corrected for multiple comparisons in time using a cluster-based permutation procedure described below.

We used a whole-brain analysis to identify brain regions showing consistent differences between cued and uncued high frequency power across participants. With iEEG, the precise placement of electrodes is different for each participant, which limits our ability to examine spatially resolved effects across participants. We overcame this limitation by spatially smoothing electrode effects so they could be projected onto a low-density ($1\text{ cm} \times 1\text{ cm}$) mesh covering the cortical surface of an MNI N27 standard brain¹². For each participant, we averaged z-scored high-frequency power during the 1-second preparatory phase from all electrodes that were within 12.5 mm of a given mesh node (Supplementary Fig. 2). Only mesh nodes that contained electrodes from three or more participants were evaluated in a paired t-test across participants. We corrected for multiple comparisons in space using a cluster-based permutation procedure. This analysis was repeated for the 1-second encoding phase. Whole-brain t-maps (Fig. 1d-e) were rendered on the vertices of a standard brain by computing the average value of all mesh nodes that were within 12.5 mm of that vertex (3-D Gaussian kernel, SD of 4.2 mm). Dark gray regions contain data from fewer than three participants.

Intracranial EEG: Attention-Related Responses

After using the whole-brain analysis and time-series analysis described above to identify the locations and timing of attention-related responses, we measured the trial-by-trial magnitude of the responses in each region to determine whether the regions' responses modulated memory accuracy, and/or were correlated with one another. We defined a 500 ms window that captured the time period of significant across-participant attention dynamics in each region: this was -500 to 0 ms preceding word onset for the anterior temporal lobe, and 250 to 750 ms following word onset for the posterior temporal and frontal cortices. We defined the dynamic response of a given trial to be the mean value in this window, minus the mean value of the 500 ms immediately preceding and 500 ms immediately following the window. This subtraction effectively removes slow fluctuations in baseline power, often described as auto-correlations^{17,18}, so that our trial-level measure reflects the dynamic, attention-related, response observed across participants.

We evaluated whether the observed attention-related responses in each region modulated memory encoding by comparing two conditions: correctly-remembered cued words versus forgotten cued words (Fig. 1i-k). We hypothesized that if the observed signals were required for memory formation, they would be absent in cued trials that were subsequently forgotten. We required at least 5 trials per condition to include it in the analysis. Participants rarely forgot cued words in sessions that passed our strict behavioral criteria (8 ± 2 total errors per participant; 8 of 18 participants with fewer than 5), so we included sessions that did not pass our criteria for this analysis (17 ± 5 total errors per participant; 5 of 18 participants with fewer than 5). We then tested for significant differences in the responses using a paired t-test across participants. We next evaluated whether the early response in the anterior temporal lobe on any given trial predicted the subsequent responses in the posterior temporal and frontal attention areas. For each participant, we computed the rank correlation between the trial-by-trial response in pairs of regions, using only correctly remembered cued trials. We then tested for significant across-participant correlations using a one-sample t-test of the resultant correlation coefficients.

Single Unit Recordings: Data Processing

We manually identified single units offline and used quantitative metrics of isolation quality to select units for subsequent analysis (Supplementary Fig. 11). After identifying a list of channels with potential single unit activity, we loaded the bandpass time series of each channel, one at a time, into Plexon Offline Sorter (Plexon, Inc; TX) for manual spike sorting. We converted the continuous voltage time series into a population of voltage snippets (1.067 ms long, 30 samples) that crossed a manually defined voltage threshold. We set the threshold such that random noise fluctuations in the signal would occasionally cross the threshold and be captured as a noise snippet. We projected each snippet into principle component space, and only retained isolated units that were separable from each other and separable from noise throughout the duration of the experiment. Based on these criteria we identified 623 putative single units, of which 302 passed additional quantitative criteria of isolation quality^{17,19}, mean spike-rate, and a minimum number of trials-per-condition (Supplemental Fig. 11). In order to maintain independent samples for statistical testing, we only allowed an individual unit to contribute once to the dataset, even if it was recorded on multiple days, leaving us with 197 independent units (Supplemental Fig. 12).

Among this pool of neurons we excluded any trials that appeared to be contaminated by transient noise. The method we used was analogous to the one we used to identify noisy iEEG channels and trials. In this case, we iteratively removed trials such that no individual trial had an across-unit mean z-score spike count above 2.8. We found that this threshold consistently eliminated trials that were deemed contaminated based on qualitative assessment of epileptic activity and/or transient noise by the clinical team.

Single Unit Recordings: Analysis

We used a 500 ms boxcar sliding window, with steps of 50 ms, to calculate a continuous estimate of time-evolving spike counts each trial. We then calculated the z-transformed, baseline-corrected spike count by subtracting the mean and dividing by the standard deviation of the spike count during the baseline period, defined as the 500 ms interval following the orientation cue (same baseline period as in the iEEG analysis). For each unit, the mean z-scored spike rate was computed at each time point separately for cued and uncued trials. Similarly, fanofactor was calculated for each unit and each time point as the variance divided by the mean spike count for that time point across trials. Population average spike rate and fanofactor were calculated across units, and significant differences between cue conditions were assessed at each time point using a paired t-test across units. We corrected for multiple comparisons in time using a cluster-based permutation procedure.

For each pair of recorded units, noise correlations were calculated separately for cued and uncued trials as the correlation in spike counts across trials²⁰ (Supplementary Fig. 5). This was done using a larger time window of 2.5 seconds (–1 to 1.5 s relative to word onset) to stabilize spike-count variation, though the results were consistent across a wide range of window sizes. We used a linear classifier to predict attentional state from population spiking activity during this same time window. We iteratively tested whether population activity predicted attentional state for each trial using a leave-one-out logistic regression with early stopping²¹. Each iteration we removed one trial to test the model, and randomly selected

20% of the remaining trials (equally drawn from cued and uncued) for the early stopping test. We repeated this iterative procedure 2000 times for each session using true (unshuffled), decorrelated (shuffled trial numbers but intact trial labels for each neuron), and random (trial labels randomized for each neuron on each iteration) data to estimate the mean classifier performance for each manipulation. Significant differences between manipulations were assessed using a paired t-test across sessions.

Statistics

Unless otherwise specified, significant differences were assessed using a paired two-tailed t-test between the cued and uncued conditions, with the independent sample being participants or neurons (unpaired two-tailed t-tests were used when the independent sample was trials). We corrected for multiple-comparisons in time series and brain-wide data using a non-parametric permutation procedure. Because of this non-parametric procedure, there is no underlying assumptions of normality or equal variance in the data that would require formal testing. No statistical methods were used to pre-determine sample sizes (number of participants, subjects, or trials) but our sample sizes are similar to those reported in previous publications^{12,17}. Data collection and analysis were not performed blind to the conditions of the experiment. No data was excluded except for: sessions and/or participants that did not meet our behavioral criterion, electrodes, neurons and trials that were excessively noisy, and units that were not well isolated. The data from each of our three analyses (human iEEG, single-unit, pre/post resection) are extremely rare and were collected over the course of 3 years. Thus we have not replicated the experiments reported here in another cohort of participants.

Statistics: Cluster-based Permutation Procedure

We corrected for multiple comparisons in space or time using a non-parametric cluster-based permutation procedure to control family-wise error^{12,22}. For whole-brain analysis, we corrected for multiple comparisons at each of the mesh nodes by first identifying clusters of neighboring mesh nodes that showed a significant effect in the same direction (positive or negative) and were within 12.5 mm of each other. For time series analysis across trials, units, or participants, we corrected for multiple comparisons at each time point by first identifying clusters of adjacent time points that showed a significant effect in the same direction (positive or negative). For each cluster (whether spatial or temporal), we computed a cluster statistic as the absolute value of the sum of the t-statistics for all elements within that cluster. In this way, a large cluster statistic can arise from a wide-spread but moderate effect, or a localized but strong effect. Finally, we created a vector-valued cluster statistic from the ranked list of cluster statistics (where the length of the vector equals the number of identified clusters). This “true” vector-valued cluster statistic must be compared to an empiric distribution of such statistics to determine whether any of the observed effects would be expected by chance.

We created an empiric distribution of vector-valued cluster statistics by randomly reassigning labels (cued vs uncued) at the level of participants, units, or trials (for example unit and electrodes), then computing a synthetic vector-valued cluster statistic for each of 5000 permutations. In each permutation, we matched the synthetic vector’s length to that of

the true vector. In permutations that had more clusters than the true data, the lower-ranked cluster statistics were not included in the vector, and for permutations with fewer clusters than the true data, the resultant vector was padded with zeros. The empiric distribution of vector-valued cluster statistics was used to calculate the likelihood of observing each cluster in the true data. For instance, we calculated a p-value for the second-biggest observed cluster statistic by counting the number of second-biggest cluster statistics that were greater or equal to the true statistic, divided by the total number of permutations. This method of comparing each cluster to its rank-matched surrogate requires an additional correction for multiple comparisons to maintain the familywise error rate²². We therefore used the false discovery rate²³ to identify clusters that remained significant (two-tailed $p < 0.05$) after correction for multiple comparisons.

Data Availability

The data that support the findings of this study are available from the corresponding author upon request and are also available for public download at <https://neuroscience.nih.gov/ninds/zaghloul/downloads.html>.

Code Availability

Custom code used to generate the findings of this study are available upon request.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Timothy Sheehan, Michael Trotta, Alex Vaz, Rafi Haque, Andrew Yang, Sri Damera, Baltazar Zavala, Rachel Ellenbogen, and Tora Bonnevie for assistance with data collection. We thank Bruno Averbeck, David Leopold, Alex Martin, Baltazar Zavala, and Tora Bonnevie for helpful discussions. We thank Rob Franklin and Nick Halper for their technical support with collecting unit data using the Cereplex I from Blackrock Microsystems, Inc. This work was supported by the Intramural Research Program of the National Institute for Neurological Disorders and Stroke. We are indebted to all participants who have selflessly volunteered their time to participate in this study.

References

1. Chun MM, Golomb JD, Turk-Browne NB. A taxonomy of external and internal attention. *Annual Review of Psychology*. 2011; 62:73–101.
2. Cohen M, Maunsell J. Attention improves performance primarily by reducing interneuronal correlations. *Nature Neuroscience*. 2009; 12:1594–1600. [PubMed: 19915566]
3. Baldauf D, Desimone R. Neural Mechanisms of Object-Based Attention. *Science*. 2014; 344:424–427. URL <http://science.sciencemag.org/content/344/6182/424>. [PubMed: 24763592]
4. Hopfinger JB, Buonocore MH, Mangun GR. The neural mechanisms of top-down attentional control. *Nature Neuroscience*. 2000; 3:284. [PubMed: 10700262]
5. Maunsell JHR. Neuronal Mechanisms of Visual Attention. *Annual Review of Vision Science*. 2015; 1:373–391.
6. Ruff DA, Cohen MR. Attention can either increase or decrease spike count correlations in visual cortex. *Nature Neuroscience*. 2014; 17:1591–1597. [PubMed: 25306550]
7. Kahana, MJ. *Foundations of Human Memory*. Oxford University Press; New York: 2014. reprint edition edn

8. Gazzaley A, Nobre AC. Top-down modulation: bridging selective attention and working memory. *Trends Cogn Sci.* 2012; 16:129–35. [PubMed: 22209601]
9. Zar, JH. *Biostatistical analysis.* 5. Pearson Education Inc; 2010.
10. Brown P, et al. Dopamine dependency of oscillations between subthalamic nucleus and pallidum in Parkinson's disease. *J Neurosci.* 2001; 21:1033. [PubMed: 11157088]
11. Hautus MJ. Corrections for extreme proportions and their biasing effects on estimated values of d . *Behavior Research Methods, Instruments, & Computers.* 1995; 27:46–51. URL <https://link.springer.com/article/10.3758/BF03203619>.
12. Haque R, Wittig J Jr, Damera S, Inati SK, Zaghoul KA. Cortical low-frequency power and progressive phase synchrony precede successful memory encoding. *Journal of Neuroscience.* 2015
13. Trotta MS, et al. Surface based electrode localization and standardized regions of interest for intracranial eeg. *Human brain mapping.* 2017
14. Oosterhof NN, Wiestler T, Downing PE, Diedrichsen J. A comparison of volume-based and surface-based multi-voxel pattern analysis. *Neuroimage.* 2011; 56:593–600. [PubMed: 20621701]
15. Burke JF, Ramayya AG, Kahana MJ. Human intracranial high-frequency activity during memory processing: neural oscillations or stochastic volatility? *Current Opinion in Neurobiology.* 2015; 31:104–110. [PubMed: 25279772]
16. Manning JR, Jacobs J, Fried I, Kahana MJ. Broadband shifts in LFP power spectra are correlated with single-neuron spiking in humans. *J Neurosci.* 2009; 29:13613–13620. [PubMed: 19864573]
17. Jang AI, Wittig JH, Inati SK, Zaghoul KA. Human Cortical Neurons in the Anterior Temporal Lobe Reinstatement Spiking Activity during Verbal Memory Retrieval. *Current Biology.* 2017; 27:1700–1705.e5. URL [http://www.cell.com/current-biology/abstract/S0960-9822\(17\)30550-X](http://www.cell.com/current-biology/abstract/S0960-9822(17)30550-X). [PubMed: 28552361]
18. Yaffe RB, et al. Reinstatement of distributed cortical oscillations occurs with precise spatiotemporal dynamics during successful memory retrieval. *Proceedings of the National Academy of Sciences.* 2014; 111:18727–18732.
19. Joshua M, Elias S, Levine O, Bergman H. Quantifying the isolation quality of extracellularly recorded action potentials. *Journal of Neuroscience Methods.* 2007; 163:267–282. [PubMed: 17477972]
20. Averbeck BB, Latham PE, Pouget A. Neural correlations, population coding and computation. *Nature Reviews. Neuroscience.* 2006; 7:358–366. [PubMed: 16760916]
21. Fukushima M, Saunders RC, Leopold DA, Mishkin M, Averbeck BB. Differential coding of conspecific vocalizations in the ventral auditory cortical stream. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience.* 2014; 34:4665–4676. [PubMed: 24672012]
22. Maris E, Oostenveld R. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods.* 2007; 164:177–190. [PubMed: 17517438]
23. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: a practical and powerful approach to multiple testing. *Journal of Royal Statistical Society, Series B.* 1995; 57:289–300.

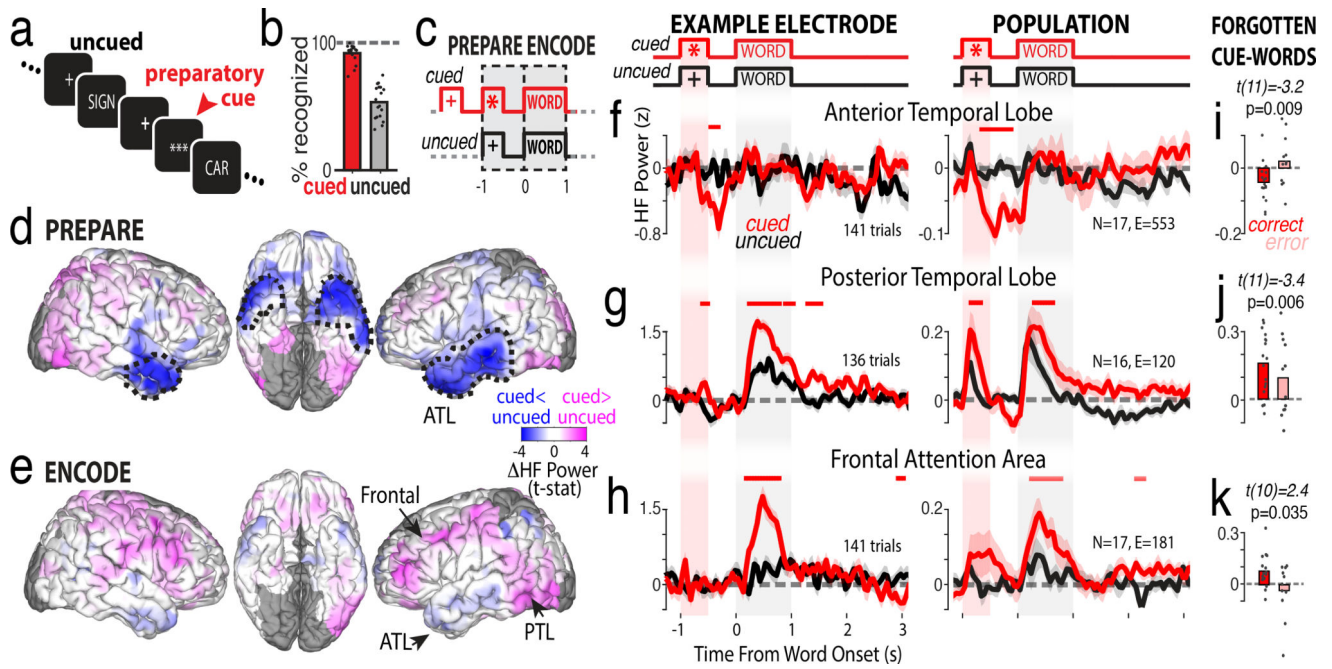


Figure 1.

Preparatory suppression of anterior temporal lobe improves memory encoding. a-c) Behavioral task that dissociates memory and attention. b) Mean recognition performance ($n=18$ participants). d,e) Whole-brain cross-participant differences in iEEG high-frequency power for cued versus uncued words during the preparation and encoding ($n=18$ participants). Dashed lines indicate regions of significant differences, corrected. f-h) Time-resolved high-frequency power for cued (red) versus uncued (black) words in individual example electrodes (left, locations marked as arrows in e, trial count is sum of cued and uncued) and across participants (right, N =number of participants, E =total number of electrodes from that region). Shaded region indicates s.e.m. across trials (left) or participants (right). Red line indicates significant differences from a two-tailed t-test, corrected. i-k) Cross-participant comparison of attention-related response magnitude for correctly remembered vs forgotten cued words (paired t-test, two tailed, uncorrected).

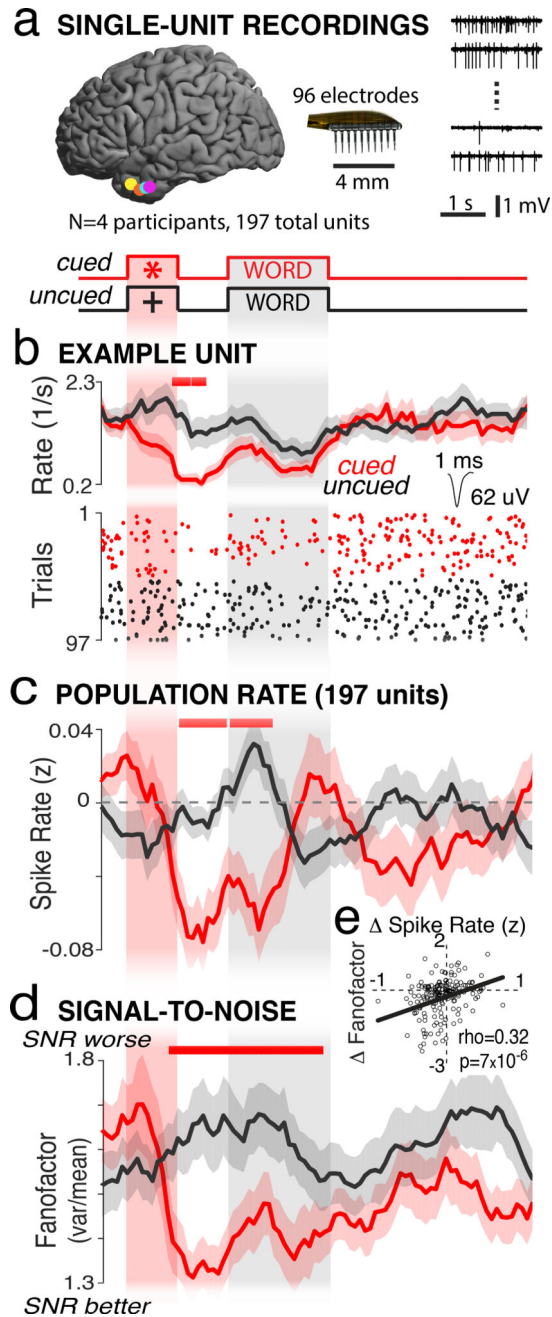


Figure 2.

Anterior temporal lobe suppression improves signal-to-noise in spiking neurons. a) Location of microelectrode arrays in four participants (left) and example time series from four microelectrodes (right). b) Raster plot (bottom) and PSTH (top) from an example unit showing preparatory suppression. c) Population spike rate and d) fanofactor show preparatory suppression. Shaded regions indicate mean \pm s.e.m. across trials (b; n=97) or units (c,d; n=197). Red line (b,c,d) indicates significant differences, corrected. e) Spearman's rank correlation between the preparatory change (–500 to 0 ms) in fanofactor and spike rate. Each point is a single unit.

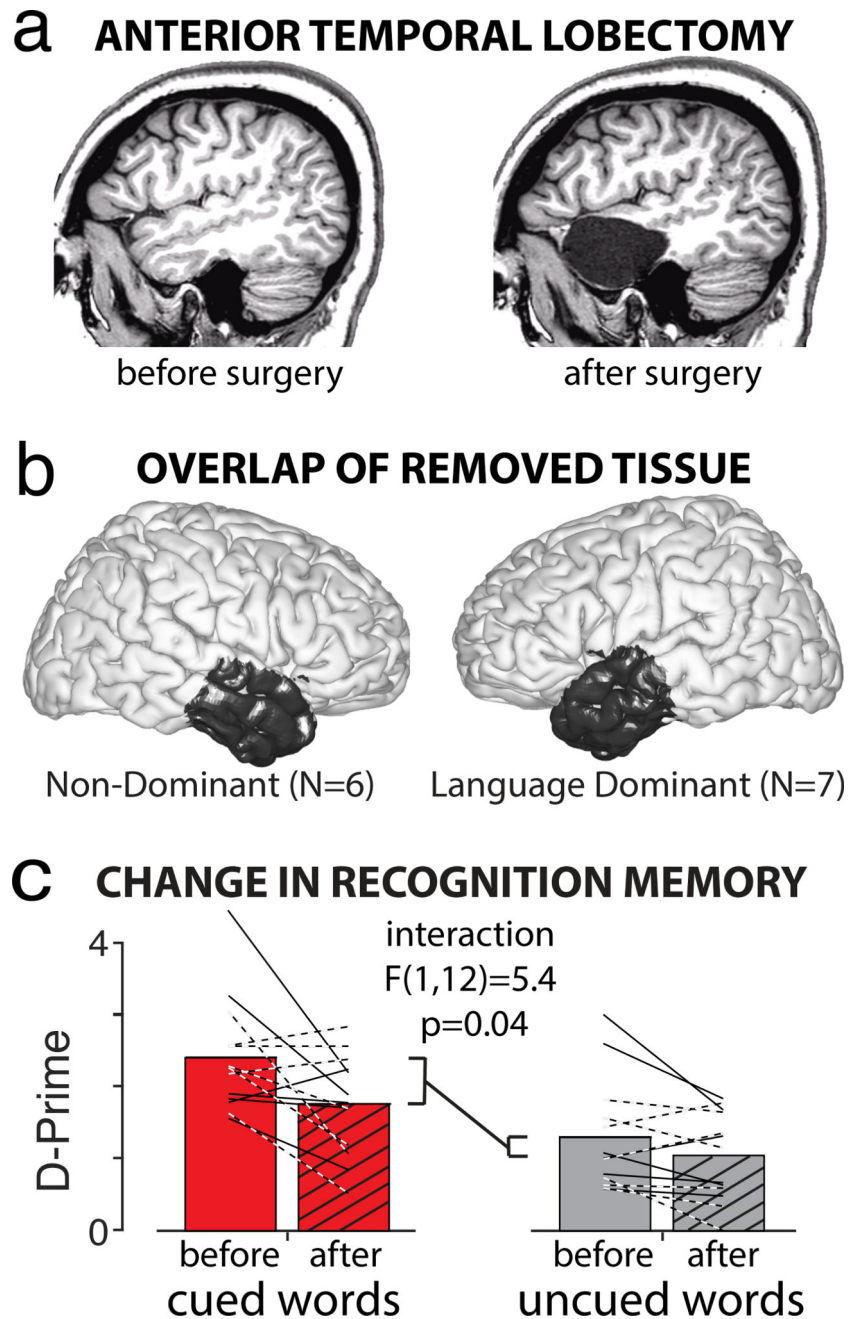


Figure 3. Anterior temporal lobe is required for attention-enhanced memorization. a) Anatomical MRIs from a participant before and after surgical removal of the anterior temporal lobe. b) Overlap of removed tissue across 13 participants. c) Recognition accuracy (D-Prime) for cued words was significantly more impaired than for uncued words 3-months after surgery (two-way, repeated measures ANOVA, with surgery and attention condition being the factors; $n=13$). Dashed lines indicate participants with a language-dominant resection.