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. Zaghloul1*

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 electroencephalography signals implicate this region as a source of executive control for attentional selection. Consistent with this interpretation, its surgical removal causes 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 the signal-to-noise ratio2. In contrast, selection processes occur in frontoparietal 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

---

1Surgical Neurology Branch, NINDS, National Institutes of Health, Bethesda, MD, USA. 2Office of the Clinical Director, NINDS, National Institutes of Health, Bethesda, MD, USA. *e-mail: kareem.zaghloul@nih.gov

---

Fig. 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 across-participant differences in intracranial electroencephalograms (EEG) high-frequency power for cued versus uncued words during the preparation and encoding phases (n = 18 participants). Dashed lines indicate regions of significant differences, corrected. Largest cluster: P = 0.0005. 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 regions indicate s.e.m. across trials (left) or participants (right). Red lines indicate significant differences from a two-tailed t test, corrected. Largest cluster, P = 0.0005. i–k, Across-participant comparison of attention-related response magnitude for correctly remembered vs. forgotten cued words (ATL: t11 = −3.2, P = 0.009; posterior temporal lobe (PTL): t11 = 3.4, P = 0.006; frontal attention area: t10 = 2.4, P = 0.035; paired two-tailed t test, uncorrected).
selection processes can be found in brain areas that respond only to attentional cues.

To identify attention mechanisms that enhance memory, we recorded intracranial electroencephalograms in 18 epileptic neurosurgery patients as they memorized words that were cued by a row of asterisks (preparatory cue; Fig. 1a.b and Supplementary Fig. 1). Cued words were remembered significantly more often than uncued words (a χ² test was evaluated for each session; only sessions for which \( P < 0.05 \) were included). To isolate attention signals, we contrasted high-frequency power (70–200 Hz) for cued words and example time series from four microelectrodes (right). Fano factor (variation over mean) and spike rate (per s) show preparatory suppression. Shaded regions indicate significant differences, corrected. SNR, signal-to-noise ratio. 

Fig. 2 | ATL suppression improves signal-to-noise results 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 peristimulus time histogram (top) from an example unit showing preparatory suppression. c, d, Population spike rate (c) and Fano factor (d) show preparatory suppression. Shaded regions indicate mean ± s.e.m. across trials (b: \( n = 97 \)) or units (c, d: \( n = 197 \)). Red lines (b–d) indicates significant differences, corrected. SNR, signal-to-noise ratio. e, Spearman’s rank correlation between the preparatory change (−500 to 0 ms) in Fano factor and spike rate (\( p = 0.32; P = 7 \times 10^{-6} \)). Each point is a single unit.

The time course of high-frequency power changes in the ATL revealed a robust and significant decrease following the preparatory cue in individual electrodes and across the population of participants (Fig. 1f). For uncued words, there was no change in high-frequency power before, during, or after word presentation (Fig. 1f), 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 or in 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). However, power was significantly greater for cued words relative to uncued words. This response illustrates the modulatory effects of attention in visual processing preceding a to-be-remembered word and the encoding phase during which the word is read, interpreted, and memorized (Fig. 1c).

During the preparation phase, high-frequency power in the anterior temporal lobe (ATL) was significantly lower for cued words relative to uncued words (Fig. 1d and Supplementary Figs. 2 and 3). In contrast, during the encoding phase, the ATL 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 and frontal cortex.

Fig. 3 | ATL is required for attention-enhanced memorization. a, Anatomical MRIs from a participant before and after surgical removal of the ATL. 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 as factors; \( n = 13 \), interaction \( F_{1,12} = 5.4, P = 0.04 \)). Dashed lines indicate participants with a language-dominant resection.
and working memory\(^1\). 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. 1b). The lack of response to uncued words, combined with large responses to cued words, was consistent with this region’s role in attentional selection\(^1\). For both regions, the attention-related response was significantly weaker for cued words that were subsequently forgotten (Fig. 1j,k), consistent with the established roles of these brain regions 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 lobe or frontal attention area. Indeed, the trial-by-trial relationships were significantly negatively correlated (P=0.0001; Supplementary Fig. 5), 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 ATL in four participants (Fig. 2a). An example neuron showed no significant change in its spike rate of 2 spikes per s when the participant viewed uncued words that were subsequently remembered (Fig. 2b). However, when the preparatory cue appeared, spike rates significantly decreased for 500 ms (Fig. 2b). The response of the example neuron shown in this figure mirrored that of the recorded population of 197 units. The population spike rate decreased significantly following the preparatory cue and remained low during the first 500 ms of the encoding period (Fig. 2c).

Visual attention improves signal-to-noise ratios in spiking neurons, measured as decreased Fano factor for individual neurons and changes in noise correlations among pairs of neurons\(^2\). We did not observe any systematic changes in noise correlations related to attention in the ATL (Supplementary Fig. 6). However, we observed a robust decrease in population-averaged Fano factor immediately following the preparatory cue (Fig. 2d). The preparatory decrease in Fano factor overlapped in time with the preparatory decrease in spike rate, suggesting that these processes were 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 Fano factor during the 500 ms preceding word onset (Fig. 2e). This suggests that the widespread suppression of high-frequency power across the anterior temporal lobe measured by iEEG reflects both the suppression of neural spiking activity and improved signal-to-noise ratios.

As a causal test of ATL 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 ATL 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 and nondominant temporal lobectomies (Supplementary Fig. 7). These results implicate a critical role for preparatory suppression of the ATL: making the semantic concepts and meanings of our experiences more salient for improved memory formation.

**Methods**

Methods, including statements of data availability and any associated accession codes and references, are available at https://doi.org/10.1038/s41593-018-0148-7.

Received: 19 January 2018; Accepted: 4 April 2018; Published online: 21 May 2018

**References**

1. Chun, M. M., Golomb, J. D. & Turk-Browne, N. B. *Annu. Rev. Psychol.* **62**, 73–101 (2011).
2. Cohen, M. R. & Maunsell, J. H. *Nat. Neurosci.* **12**, 1594–1600 (2009).
3. Baldauf, D. & Desimone, R. *Science* **344**, 424–427 (2014).
4. Hopfinger, J. B., Buonocore, M. H. & Mangun, G. R. *Nat. Neurosci.* **3**, 284–291 (2000).
5. Maunsell, J. H. R. *Annu. Rev. Vis. Sci.* **1**, 373–391 (2015).
6. Ruff, D. A. & Cohen, M. R. *Nat. Neurosci.* **17**, 1591–1597 (2014).

**Acknowledgements**

We thank T. Sheehan, M. Trotta, A Vaz, R. Haque, A. Yang, S. Damera, B. Zavala, R. Ellenbogen, and T. Bonnevie for assistance with data collection. We thank B. Averbeck, D. Leopold, A. Martin, B. Zavala, and T. Bonnevie for helpful discussions. We thank R. Franklin and N. 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.

**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 interests**

The authors declare no competing financial interests.

**Additional information**

Supplementary information is available for this paper at https://doi.org/10.1038/s41593-018-0148-7.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to K.A.Z.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
Methods

Participants. Eighteen neurosurgery patients (5 male; age 35 ± 3 years; FSIQ 99 ± 5; mean ± s.e.m.; see Supplementary Table 1 and Reporting Summary) with drug-resistant epilepsy underwent a surgical procedure in which platinum electrodes were implanted beneath 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 of the 18 participants additionally and specifically consented to have microelectrode 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 rooms during their stays 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, and all microelectrode arrays were at least 2 cm from the closest contact that detected ictal or interictal activity. Second, we removed any data from grooming activities. Patients in the postictal phase were examined. To assess the impact of the recording electrode placement on the results, we compared the recognition accuracy for uncued words was significantly better than for uncued words and (ii) recognition performance on the recall test indicated that they followed our instructions with and without attentional engagement.

Successful memorization of cued and uncued words in our task is likely supported by episodic memory processes of familiarity and recollection\(^1\), though it is possible that short-term memory processes such as working memory could be involved.\(^2\) Working memory could be used to subvocally 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 s, depending on list length, which is longer than the intervals commonly used for working memory tests\(^2\). 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 subvocally rehearse the words. Third, performance on the recall test indicated that they followed our instructions to specifically memorize the cued words, which suggests that participants were not expending effort to actively rehearse uncued words that were successfully recognized.

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 during 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 isoperformance data across participants who showed a wide range of natural aptitude for the task.

Effect of surgery. Only participants who passed our performance criteria on a seen/unseen recognition test during preoperative testing were included in our analysis of the behavioral effects of surgery. For each participant in this analysis, task parameters (for example, list length) and antiepileptic pharmaceutical treatments were identical in the preoperative and postoperative test sessions. Preoperative testing sessions were preceded by one to three training sessions used to identify ideal task parameters. Postoperative testing sessions were preceded by approximately 15 min of training to refamiliarize the participant with the task. We quantified behavioral performance in these pre- and postoperative sessions using D-prime. This measure of recognition accuracy incorporates false-alarm rates to account for an individual’s bias toward 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 to the numerator and 1.0 to the denominator of the proportion of seen trials to 0.5. D-prime was computed from each participant’s test data by aggregating data across postoperative 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 1,000 Hz using a Nihon Kohden EEG data acquisition system. Subdural contacts were arranged in both grid and strip configurations with an intercontact spacing of 5 or 10 mm. Electrode localization was accomplished by co-registering the postoperative CTs with the postoperative MRIs using both FSL Brain Extraction Tools and FLIRT image processing packages. Preoperative EEGs after postoperative 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\(^12\).

Our first step in cleaning the iEEG signals was to eliminate any electrodes identified by the clinical team 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 (electrical activity, physical movement of the participant, or external sources of transient electrical perturbations) using a procedure we adapted from the EEG analysis software package FieldTrip\(^13\). For each channel and trial, we computed the variance of the voltage trace during the 5.25 s surrounding the encoding period (−1.5 s before word onset to 3.75 s 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 + 1.5 × interquartile range, where w_quart is a user-specified parameter. If any trials exceeded the threshold, the maximum variance trials were iteratively removed, and the quartiles were recalculated at each step. If any of the removed trials caused a change in the maximum thresholds that exceeded the threshold, that trial or channel was removed. 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_quart of 0.5 (equivalent to the mean) worked well in catching extreme values of noisy data from the epileptologist on our clinical team. After systematically removing noisy channels and trials, we removed the effects of a common reference by subtracting a global common average, computed across all channels separately for each trial and timepoint. We confirmed that any potential noise artifacts caused by saccadic eye movements could be accounted for using the attention-related effects reported in the anterior temporal lobe (Supplementary Fig. 9).

Power analysis. We estimated a continuous time measure of high-frequency power using 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 activity\(^14\). We log-transformed the power estimates\(^15\), downsampled in time using a sliding 100 ms boxcar window with a step of 50 ms, baseline-corrected and z-transformed each timepoint by subtracting the mean and dividing by the s.d. 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 timepoint. For population-average electrode effects in predefined regions of interest (Suppl. Fig. 9), we first calculated a mean cued and mean uncued time series for each electrode, averaged the time series from electrodes in that region for each participant, and then calculated a paired t test across participants at every timepoint. For individual electrodes and population
averages, we corrected for multiple comparisons in time using a cluster-based permutation procedure, described below.

We used whole-brain analysis to identify brain regions showing consistent differences between cued and uncued high-frequency activity 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 cm × 1 cm) mesh covering the cortical surface of an MNI N27 standard brain16. For each participant, we averaged z-scored high-frequency power during the 1-s 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-s 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, s.d. of 4.2 mm). Dark gray regions contain data from fewer than three participants.

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 with significant across-participant effects (Supplementary Fig. 3). For each participant, this window was determined by visual inspection. For 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 power, power that would occasionally cross the threshold and be captured as a noise snippet. Voltage thresholds were set such that random noise fluctuations in the average signal would occasionally cross the threshold and be captured as noise 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 baseline activity 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 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 quality17,19, mean spike rate, and a minimum number of trials per condition (Supplementary Fig. 4). Thus we have not replicated the experiments reported here in another cohort of 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–l). 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 missed 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 posterior temporal lobe on any given trial predicted the subsequent responses in the posterior temporal and frontal attentional 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 recording. Data processing. We manually identified single units offline and used quantitative metrics of isolation quality to select units for subsequent analysis. After identifying a list of channels with potential single-unit activity, we loaded the globally referenced and passband (0.3 to 3 kHz) time series of each channel, one at a time, into Plexon Offline Sorter (Plexon, Inc.; TX) for manual spike sorting2. 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 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 quality17,19, mean spike rate, and a minimum number of trials per condition (Supplementary Fig. 1). 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 (Supplementary Fig. 1h).

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. 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 for each trial. We then calculated the z-transformed, baseline-corrected spike count by subtracting the mean and dividing by the s.d. 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 neural score was computed at each timepoint separately for cued and uncued trials. Similarly, the Fano factor was calculated for each unit and each timepoint as the variance divided by the mean spike count for that timepoint across trials. Population average spike rated and Fano factor were calculated across units, and significant differences between cue conditions were assessed at each timepoint using paired t tests 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 trials (Supplementary Fig. 5). This was done using a larger time window of 2.5 s (−1 to +1.5 s relative to word onset target) across participants. We corrected for multiple comparisons in space using a cluster-based permutation procedure. This analysis was repeated for the 1-s 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, s.d. of 4.2 mm). Dark gray regions contain data from fewer than three participants.

Statistics. Unless otherwise specified, significant differences were assessed using a paired two-tailed t test between the cued and uncued conditions, with the number of independent samples being the participant. Often, when processing data from patients with less than 5 trials, we performed the analysis with fewer than 5). We then tested for significant differences in the responses using a paired t test across participants. For each iteration we removed one trial to test the model, and randomly selected the remaining 20% of the remaining trials (equally drawn from cued and uncued) for the early stopping test. We repeated this iterative procedure 2,000 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 iteself during data collection) to estimate the mean classifier performance for each manipulation. Significant differences between manipulations were assessed using a paired t tests across sessions.

Cluster-based permutation procedure. We corrected for multiple comparisons in time series and brain-wide data using a nonparametric permutation procedure. Because of this nonparametric procedure, there are no underlying assumptions of normality or equal variance in the data that would require formal testing. No statistical methods were used to predetermine sample sizes (number of participants, subjects, or trials), but our sample sizes are similar to those reported in previous publications20,21. Data collection and analysis were not performed blind to the conditions of the experiment. No data were excluded except for sessions and/or participants who 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-/postresection) 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.
error rate². We therefore used the false discovery rate²³ to identify clusters that remained significant (two-tailed $P < 0.05$) after correction for multiple comparisons.

**Reporting Summary.** Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

**Code availability.** Custom code used to generate the findings of this study is available upon reasonable request.

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

**References**

1. Zar, J.H. *Biostatistical Analysis* (Pearson Education Inc, 2010), 5th edn.
2. Brown, P. et al. *J. Neurosci.* 21, 1033–1038 (2001).
3. Kahana, M. J. *Foundations of Human Memory* (Oxford University Press, New York, 2014).
4. Gazzaley, A. & Nobre, A. C. *Trends Cogn. Sci.* 16, 129–135 (2012).
5. Hautus, M. J. *Behav. Res. Methods Instrum. Comput.* 27, 46–51 (1995).
6. Haque, R., Wittig, J. Jr., Damera, S., Inati, S. K. & Zaghloul, K. A. *J. Neurosci.* 35, 13577–13586 (2015).
7. Trotta, M. S. et al. *Hum. Brain Mapp.* 39, 709–721 (2017).
8. Oosterhof, N. N., Wiestler, T., Downing, P. E. & Diedrichsen, J. *Neuroimage* 56, 593–600 (2011).
9. Burke, J. F., Ramayya, A. G. & Kahana, M. J. *Curr. Opin. Neurobiol.* 31, 104–110 (2015).
10. Manning, J. R., Jacobs, J., Fried, I. & Kahana, M. J. *J. Neurosci.* 29, 13613–13620 (2009).
11. Jang, A. I., Wittig, J. H. Jr., Inati, S. K. & Zaghloul, K. A. *Curr. Biol.* 27, 1700–1705.e5 (2017).
12. Yaffe, R. B. et al. *Proc. Natl. Acad. Sci. USA* 111, 18727–18732 (2014).
13. Joshua, M., Elias, S., Levine, O. & Bergman, H. *J. Neurosci. Methods* 163, 267–282 (2007).
14. Averbeck, B. B., Latham, P. E. & Pouget, A. *Nat. Rev. Neurosci.* 7, 358–366 (2006).
15. Fukushima, M., Saunders, R. C., Leopold, D. A., Mishkin, M. & Averbeck, B. B. *J. Neurosci.* 34, 4663–4676 (2014).
16. Maris, E. & Oostenveld, R. *J. Neurosci. Methods* 164, 177–190 (2007).
17. Benjamini, Y. & Hochberg, Y. *J. R. Stat. Soc. B* 57, 289–300 (1995).
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

- [ ] n/a

- [ ] Confirmed

- [ ] The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement

- [ ] An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly

- [ ] The statistical test(s) used AND whether they are one- or two-sided

- [ ] Only common tests should be described solely by name; describe more complex techniques in the Methods section.

- [ ] A description of all covariates tested

- [ ] A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons

- [ ] A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)

- [ ] For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted

- [ ] Give P values as exact values whenever suitable.

- [ ] For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings

- [ ] For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes

- [ ] Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated

- [ ] Clearly defined error bars

- [ ] State explicitly what error bars represent (e.g. SD, SE, CI)

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection

- iEEG data were acquired using Nihon Khoden's EEG data acquisition software. Single unit data were acquired using Blackrock Microsystems's NeuroPort Central Suite (v 7.0.3.0)

Data analysis

- Imaging data were analyzed using both FSL Brain Extraction Tool (BET), FLIRT software packages, and a Talairach daemon for the N27 standard brain. Single units were isolated using Plexon Offline Sorter. All visualization and statistical analyses were performed using custom code in Matlab, R15a (Mathworks, Inc.). Custom code is available upon request.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data
**Policy information about availability of data**
All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on 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.

---

**Field-specific reporting**

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☑ Life sciences  ☐ Behavioural & social sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

---

**Life sciences**

**Study design**

All studies must disclose on these points even when the disclosure is negative.

| Sample size | Sample sizes were chosen in the following manner. In order to aggregate data within a given brain region across participants, we require a minimum of three participants with intracranial electrodes implanted within that brain region. Due to clinical considerations, every participant will have a different arrangement of implanted surface and depth electrodes. Given this variability and based on our previous work, we estimated that we needed at minimum 12 participants to perform a task before we can draw significant conclusions regarding changes in neural activity in any one brain region. No explicit power calculations were performed to determine sample sizes. However, based on the electrode coverage of the participants in this study (Supplementary Fig 2), we found that 18 participants were required to provide sufficient coverage over a majority of the cortical surface. Among these 18 participants exactly 4 had Utah arrays with viable data that we included in the unit analysis. We used all available data for the resection analysis. |
| Data exclusions | Data were excluded from analysis for the following reasons. Any task session in which behavior did not meet performance criteria was excluded (21/90 experimental sessions; described in Supplementary Fig. 8). Any participant that did not complete at least two experimental sessions that met performance criteria was excluded. We excluded any channel or trial that exhibited significant noise (described in Methods, Intracranial EEG: Data Processing and Single Unit Recordings: Data Processing). |
| Replication | We did not replicate these results in a separate cohort of participants. The data presented here were captured over three years from intracranial recordings in human neurosurgical patients receiving treatment for epilepsy, and are thus extremely rare. |
| Randomization | Randomization of participants was not relevant to this study and participants were therefore not allocated into separate groups. Every participant completed an experimental session in which analyses were conducted comparing neural activity between task conditions. The order of cued and uncued words was randomized each list. |
| Blinding | Blinding was not relevant to this study. All participants performed the same behavioral task, and data were analyzed for each participant who completed at least two experimental sessions. |

---

**Materials & experimental systems**

Policy information about availability of materials

n/a Involved in the study

☑ Unique materials

☑ Antibodies

☑ Eukaryotic cell lines

☑ Research animals

☐ Human research participants

---

**Human research participants**

Policy information about studies involving human research participants

| Population characteristics | Among 18 participants that contributed electrophysiological data: 5 male; age 35 ± 3 years; FSIQ 99 ± 5; mean ± s.e.m.; |
| Method                     | Included in Study |
|----------------------------|-------------------|
| ChIP-seq                   | ✗                 |
| Flow cytometry             | ✗                 |
| Magnetic resonance imaging | ✗                 |
| n/a                        |                   |