

Plasticity and language in the anaesthetized human hippocampus

<https://doi.org/10.1038/s41586-026-10448-0>

Received: 9 April 2025

Accepted: 25 March 2026

Published online: 06 May 2026

Open access

 Check for updates

Kalman A. Katlowitz¹, Eric R. Cole^{1,15}, Elizabeth A. Mickiewicz^{1,15}, Shraddha Shah^{1,15}, Melissa Franch¹, Joshua A. Adkinson¹, James L. Belanger¹, Raissa K. Mathura¹, Domokos Meszéna^{2,3,4}, Matthew McGinley⁵, William Muñoz⁶, Garrett P. Banks¹, Sydney S. Cash^{2,4}, Chih-Wei Hsu⁷, Angelique C. Paulk^{2,4}, Nicole R. Provenza^{1,8,9,10}, Andrew J. Watrous¹, Ziv Williams⁶, Alica M. Goldman¹¹, Vaishnav Krishnan^{5,8,11}, Atul Maheshwari¹¹, Sarah R. Heilbronner^{1,5,10}, Robert Kim¹², Nuttida Rungratsameetaweemana¹³, Benjamin Y. Hayden^{1,5,8,10,16} & Sameer A. Sheth^{1,5,8,10,14,16}✉

Consciousness is a fundamental component of cognition¹, but the degree to which higher-order pattern recognition relies on it remains disputed^{2,3}. Here we demonstrate the persistence of oddball discrimination, semantic processing and online prediction in individuals under general-anaesthesia-induced loss of consciousness^{4,5}. Using high-density Neuropixels microelectrodes⁶ to record both single-unit and local-field-potential neural activity in the human hippocampus while playing a series of tones to anaesthetized patients, we found that hippocampal neurons and local oscillations retained some detection of oddball tones. This effect size grew over the course of the experiment (around 10 min), demonstrating representational plasticity. A biologically plausible recurrent neural network model showed that learning and oddball representation are an emergent property of flexible tone discrimination. Moreover, when we played language stimuli, single units and local field potentials carried information about the semantic and grammatical features of natural speech, even predicting semantic information about upcoming words. Together these results indicate that in the hippocampus, which is anatomically and functionally distant from primary sensory cortices⁷, complex processing of sensory stimuli occurs even in the unconscious state.

A central question in cognitive neuroscience is the extent to which complex information processing depends on conscious awareness. Prominent theories of consciousness propose that sophisticated pattern recognition, semantic interpretation and predictive processing all require conscious access, particularly when these computations involve integration across multiple timescales or abstraction beyond immediate sensory features^{8,9}. At the same time, evidence from psychology and neuroscience suggests that substantial processing can occur outside awareness, including perceptual discrimination, statistical learning and aspects of language comprehension^{10,11}. These findings raise the possibility that neural circuits—even those distant from sensory receptors and motor effectors—may continue to encode and transform meaningful structure in sensory input even when consciousness is disrupted¹². One context in which this question can be addressed is general anaesthesia, which provides a reversible and well-characterized state of unconsciousness⁵. Although anaesthesia profoundly alters

large-scale brain dynamics and suppresses behavioural responsiveness, several studies have reported residual sensory responses in early cortical areas during anaesthetized states⁴. Here we examined whether neural correlates of higher-order processing persist during anaesthesia in the hippocampus—a region anatomically and functionally distant from primary sensory and motor systems.

Experimental design

We performed intraoperative hippocampal recordings using Neuropixels probes⁶ in seven patients undergoing anterior temporal lobectomies (Extended Data Table 1). Recordings were conducted in the anterior body after resection of the lateral temporal cortex and before resection of mesial temporal structures such as parahippocampal gyrus and amygdala (Fig. 1a–g). After a brief baseline recording, we conducted recordings during presentation of auditory stimuli

¹Department of Neurosurgery, Baylor College of Medicine, Houston, TX, USA. ²Department of Neurology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA. ³HUN-REN Research Centre for Natural Sciences, Budapest, Hungary PPCU Faculty of Information Technology and Bionics, Budapest, Hungary. ⁴Center for Neurotechnology and Neurorecovery, Department of Neurology, Mass General Brigham, Boston, MA, USA. ⁵Department of Neuroscience, Baylor College of Medicine, Houston, TX, USA. ⁶Department of Neurosurgery, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA. ⁷Department of Integrative Physiology, Baylor College of Medicine, Houston, TX, USA. ⁸Department of Electrical & Computer Engineering, Rice University, Houston, TX, USA. ⁹Department of Bioengineering, Rice University, Houston, TX, USA. ¹⁰Neuroengineering Initiative, Rice University, Houston, TX, USA. ¹¹Department of Neurology, Baylor College of Medicine, Houston, TX, USA. ¹²Department of Neurology, Cedars-Sinai Medical Center, Los Angeles, CA, USA. ¹³Department of Biomedical Engineering, Columbia University, New York, NY, USA. ¹⁴Department of Psychiatry and Behavioral Sciences, Baylor College of Medicine, Houston, TX, USA. ¹⁵These authors contributed equally: Eric R. Cole, Elizabeth A. Mickiewicz, Shraddha Shah. ¹⁶These authors jointly supervised this work: Benjamin Y. Hayden, Sameer A. Sheth. ✉e-mail: sameer.sheth@bcm.edu

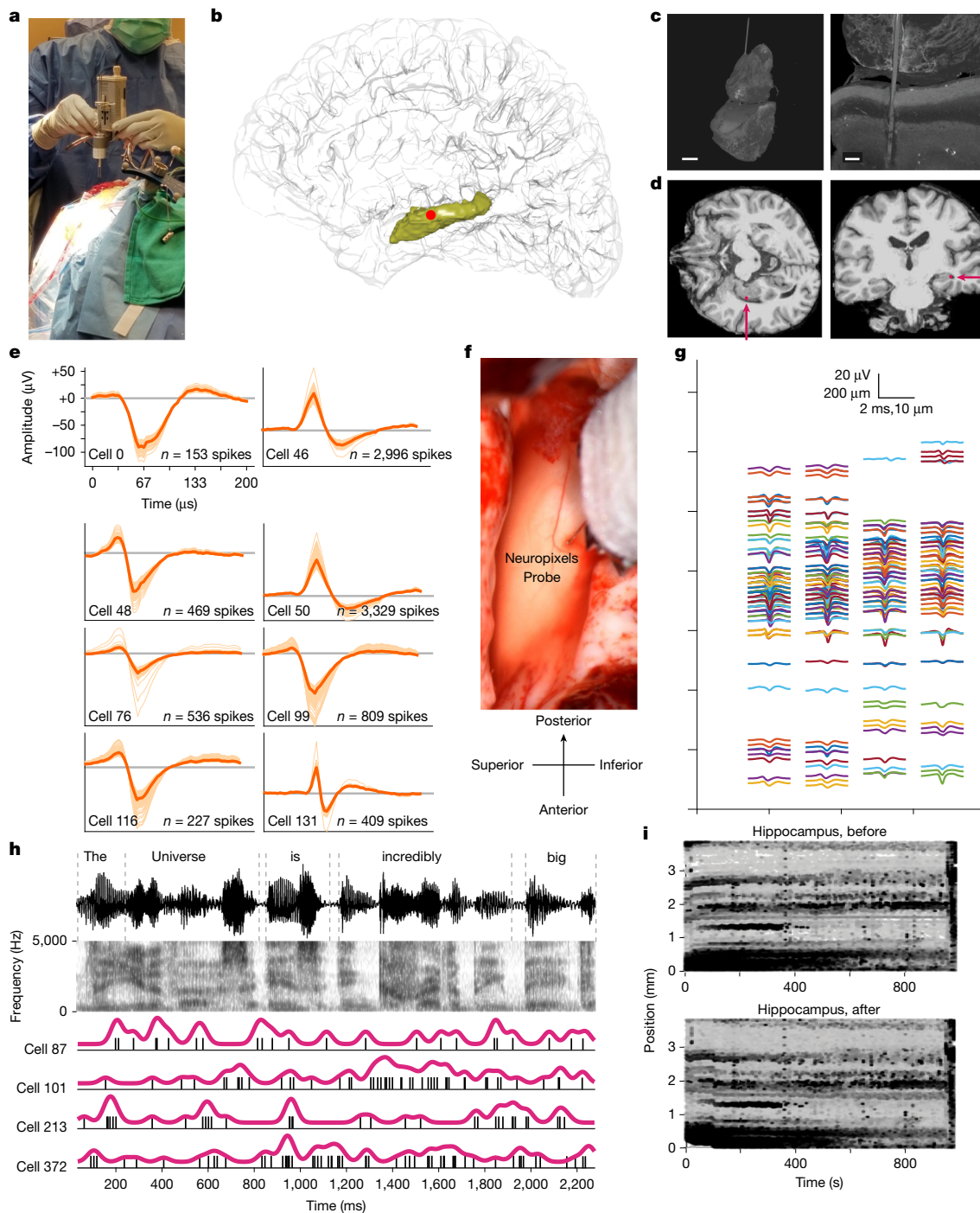


Fig. 1 | Neuropixels implantation procedure and analysis. **a**, Picture of Neuropixels insertion with robotic guidance. **b**, The probe entry site for p8 warped onto the canonical brain, illustrated with a crimson dot within the hippocampus (shown in yellow). Code from ref. 51 used to generate the image. **c**, In one patient (p8), we extracted the probe along with the tissue. Shown is a 3D reconstruction (micro-computed tomography) identifying the probe within resected hippocampal tissue (left) with coronal slice identifying the probe penetrating the hippocampus (right). The superior globule is fibrin glue adhering to the ependymal lining. **d**, Axial (left) and coronal (right) sections of T1 magnetic resonance imaging for p8. The crimson dot indicates the probe entry site, and the arrows show the trajectory of the probe. **e**, Example waveforms from units collected in our recording. The thin lines are individual spikes (n values are listed in the panel); the thick line represents the median. **f**, Photograph of the hippocampal brain tissue with the inserted Neuropixels

probe during intraoperative recording (middle right), with the anatomical orientation indicated below. **g**, Example waveforms from all units ($n = 127$) within a single hippocampal recording (p5). Each unit is represented by the waveforms at the three channels with greatest magnitude. The x axis refers to each of the four columns in the columnar arrangement of sensors. The y axis shows the position along the electrode shank. **h**, The analysis pipeline; we transcribed individual words with precise temporal alignment. Vertical lines are spike times, and magenta lines show the smoothed firing rates. **i**, Drift maps illustrating the low drift rates and high stability of Neuropixels recordings in human hippocampus. The x axis shows the time across the session. The y axis shows the depth along the electrode shank. The points (black marks) indicate individual spikes. The banding pattern reflects degree of motion before (top) and after (bottom) motion correction. For **c**, scale bars, 1,000 μm (left) and 300 μm (right).

composed of pure tones (three patients) or a podcast (four patients; Fig. 1h). Across all recordings, we isolated 651 units. Average firing rates were low (1.8 ± 1.1 Hz). Motion artefacts, a major challenge for human cortical Neuropixels recordings¹³, were markedly less conspicuous than in the cortical recordings (Fig. 1i), presumably due to the central location of the hippocampus, and because it is anchored by the dura of the middle fossa. Consistent with this hypothesis, the reduction in motion was especially clear when we compared respiratory and heartbeat frequency bands to a control recording performed in the cortex before resection ($P < 0.001$, *t*-test on power between 0.1 and 3 Hz of motion trajectories between hippocampal and cortical recordings).

Auditory monitoring during anaesthesia

In the oddball task¹⁴, three patients were presented with identical 100 ms tones interspersed with oddballs (20%, higher/lower frequency tones; Fig. 2a and Methods) with random stimulus onset asynchrony (1–3 s). Most units ($n = 122$ out of 172, 70.9%, signed-rank test, $\alpha = 0.05$) showed tone-evoked responses (Fig. 2b), consistent with established auditory responses in hippocampus¹⁵. Across all units, response latencies showed a biphasic temporal dynamic (Gaussian mixture model fit through expectation maximization; Fig. 2b,c). Units encoded tone identity ($P < 0.001$, generalized linear mixed effects (GLME); $n = 39$ out of 172, 22.7% of units, rank-sum test, $\alpha = 0.05$; Fig. 2d and Methods). There was no difference in the proportion of neurons that significantly responded to each oddball type ($P = 0.184$; $n = 172$ units; χ^2 goodness of fit test); thus, tone discrimination effects reflect a balanced shift in the population response rather than a preferred acoustic feature of one of the tones.

We next examined the representation of stimulus features. For two patients, we balanced tone identity and oddball status (Fig. 2a; $n = 150$ units). At the single-unit (Fig. 2e) and population (Fig. 2f) levels, neuronal responses differentiated standard from oddball tones ($P < 0.0001$, GLME). This divergence was most notable within the first 300 ms (24.7%, $n = 37$ out of 150 of units, signalled oddballs). Subsequent analyses focused on this epoch segment. Note that we use a symmetric non-causal Gaussian filter, which equally weighs past and future timepoints; the small amount of responding before the tone is an artefact of this filter. Moreover, because excitatory and inhibitory tone-evoked responses tended to cancel each other out, there is no visible peak in the response for the standard tone. Local field potentials (LFPs) also showed oddball-evoked responses, demonstrating a negative deflection in the evoked response potential (ERP; Fig. 2g) and an increase in gamma amplitude (Fig. 2h).

We modelled evoked responses for all units as a function of tone identity, context (standard versus oddball), and their interaction. We observed comparable encoding for all terms: tone encoding ($P < 0.0001$, GLME; 29.3% of units significant), oddball encoding ($P < 0.0001$, GLME; 24.7% of units significant) and interaction ($P < 0.01$, GLME; 22.7% of units significant). The absolute values of the betas for the oddball term were greater than the corresponding tone and mixed selectivity terms (paired *t*-test on absolute values, $P < 0.0001$ for both; Fig. 2i). Similar proportions of units showed a significant oddball effect ($n = 43$) in the two patients (37 out of 127, 29.1%; and 6 out of 23, 26.1%; $P = 0.8$, χ^2 test). The mean broadband LFP power and gamma band amplitude also demonstrated tone, oddball and mixed selectivity at similar rates across channels but reduced significance (broadband LFP: $P = 0.224$, $P = 0.038$ and $P = 0.0186$, linear mixed-effects (LME), 40.9%, 47.2% and 46.0% of channels; gamma: $P = 6.18 \times 10^{-4}$, $P = 0.024$ and $P = 2.80 \times 10^{-4}$, LME, 20.1%, 17.6% and 18.7% of channels, respectively). Tone detection effects remain robust at lower significance levels of $P < 0.01$ (55.2% of neurons) and $P < 0.001$ (28.5%), as do tone-selective effects ($P < 0.01$, 12.2%; $P < 0.001$, 4.1%). Modelled oddball detection effects remain robust at $P < 0.01$ (14.7%)

and $P < 0.001$ (9.3%), and tone–oddball interaction effects remained robust at $P < 0.01$ (16.7%) and $P < 0.001$ (10.0%).

We used a tenfold cross-validated support vector machine (SVM) to decode stimulus features on a trial-by-trial level across the population. Tone identity was robustly represented in both patients across units and LFP, with accuracy ranging between 0.61 and 0.70 ($P < 0.001$ for all, *t*-test; Fig. 2j). Oddball identity was decodable at above chance for the two patients combined ($P < 0.05$, *t*-test), albeit not in p6 alone. We found that tone identity could be decoded across all tested frequency bands (range, 0.52 to 0.79; $P < 0.001$, *t*-test). Oddball identity was decodable in some frequency bands for both patients ($P < 0.001$ for delta and high gamma bands for p5; all bands except delta and theta for p6, Fig. 2k). Thus, tone and oddball information is present in both single unit and LFP, with oddball signals having substantially weaker strength than tone identity.

High gamma phase-amplitude coupling reliably detected tone and oddball stimuli: 26.6% of channels demonstrated significant tone encoding, 21.96% of channels demonstrated significant oddball encoding and 24.07% of channels demonstrated an interaction ($P < 0.05$). Using SVM decoding, high gamma phase-amplitude coupling predicted tone identity (p5, tone accuracy = 0.553; p6, tone accuracy = 0.542), but could not decode oddballs. Low gamma phase-amplitude coupling did not significantly encode tones, oddballs or their interaction. Low gamma phase-amplitude coupling could weakly decode both tone identity and oddball status in p6 (tone accuracy = 0.505, $P < 0.001$; oddball accuracy = 0.527, $P < 0.005$) but not p5 ($P > 0.05$ for both tone and oddball).

Plasticity in the unconscious state

We next examined the temporal evolution of the oddball response. In oddball-selective units ($n = 43$), we found that oddball response grew more distinct (as inferred from decodability) over the course of the experiment (around 10 min; an example unit is shown in Fig. 3a). Splitting our task into first and second halves (for each block), we found a significant increase in oddball encoding for both patients (p5, $P = 0.01$; p6, $P < 0.001$, *t*-test; Fig. 3b). We also observed a decrease in tone identity encoding from the first to the second half of each block, raising the possibility of compensatory mechanisms ($P < 0.001$, *t*-test; Fig. 3c). Using a 50-stimulus sliding window, we found a continuous increase in oddball decoding accuracy across the 10 min duration of the experiment ($P < 0.001$, Pearson's correlation; Fig. 3d (green)). This increase in oddball performance was accompanied by a decrease in tone encoding ($P < 0.0001$; Fig. 3d (purple))¹⁶. There was a significant negative correlation between the evolution of trajectories of oddball and tone decoding accuracy ($r = -0.23$, $P < 0.03$, Pearson's test), consistent with the hypothesis that the neural population was sacrificing tone responses for the sake of oddball representations over the course of the experiment¹⁷. Phase-amplitude coupling demonstrated a significant difference in tone decoding accuracy between first half and second half of the oddball experiment (*t*-test; $P < 0.0001$) in both patients. Oddball decoding only significantly increased in the second half of the oddball task for p6 ($P < 0.0001$), not p5. Spike-frequency coupling showed a significant increase in both tone and oddball decoding between task halves, and in both patients ($P < 0.0001$).

We created neural vectors of the average standard tone as well as each individual oddball trial (43-dimensional vectors composed of the mean response of the oddball units). We found a gradual divergence in Euclidean distance between standard and oddball vectors over the course of the session ($r = 0.34$, $P < 0.0001$, LME; Fig. 3e (left)). Discriminability was even stronger when considering cosine angle, indicating that the effect is not merely a consequence of a response gain in oddball cells ($r = 0.5$, $P < 0.0001$, LME; Fig. 3e (right)). These effects were mostly consistent for individual patients (p5 distance, $r = 0.25$, $P = 0.056$; angle, $r = 0.43$, $P = 0.002$; p6 distance, $r = 0.32$, $P = 0.012$; angle, $r = 0.48$, $P = 0.0002$).

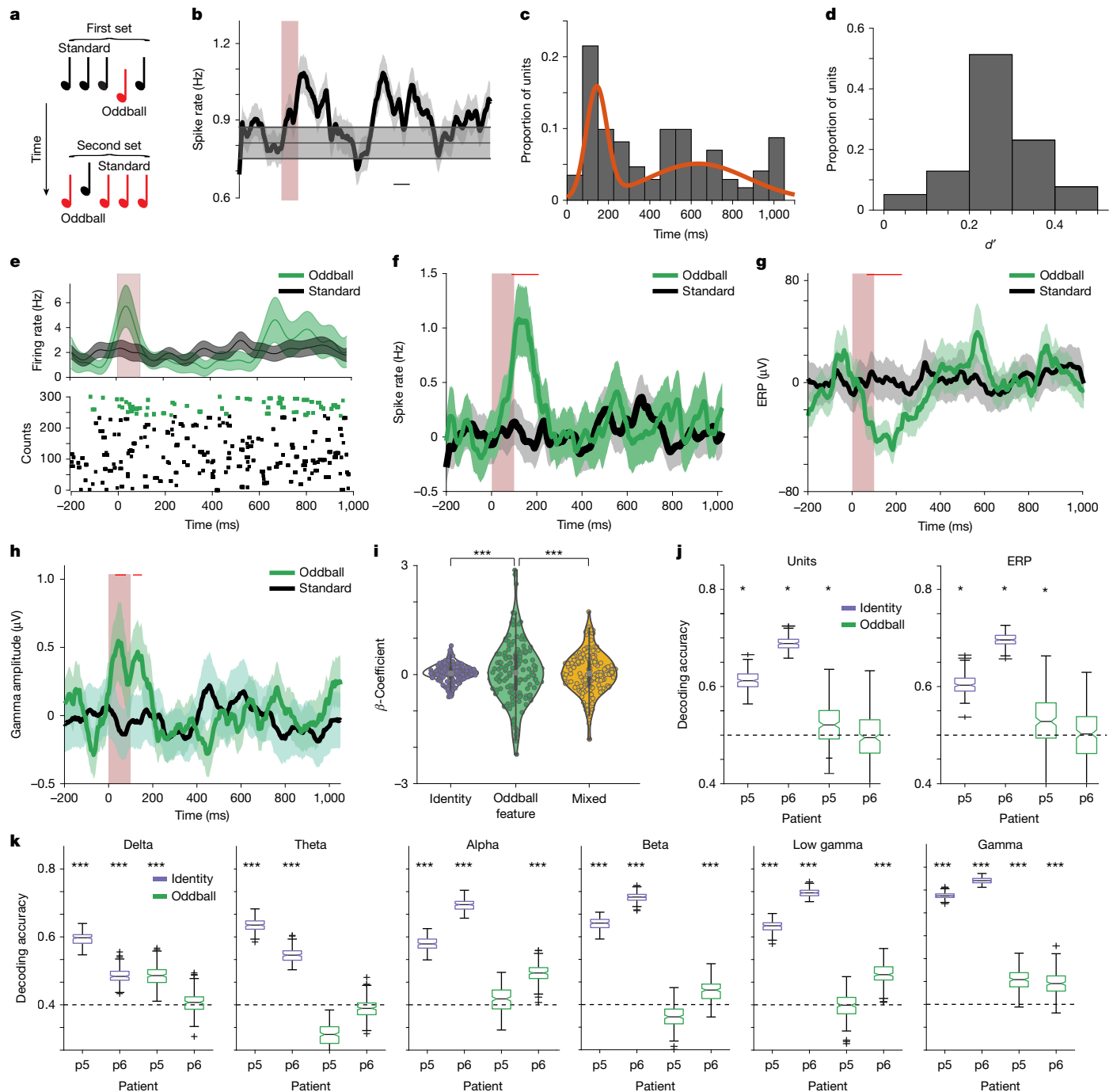


Fig. 2 | Oddball responses in the anaesthetized human hippocampus.

a, Schematic of the pure tone auditory oddball task. The oddball and standard were interchanged after 150 tones. **b**, The mean tone response ($n = 172$ units, 3 patients). The red bar indicates tone presentation (100 ms). The grey bar represents the baseline \pm s.e.m. Scale bar, 100 ms. **c**, The distribution of the tone response onset latencies. A mixed Gaussian model fit is shown in orange. **d**, d' values for units selective for tone identity ($n = 39$ units). **e**, Example unit selective for oddballs. Top, the mean \pm s.e.m. response (spike rate, Hz) to oddball (green) and standard (black). The red bar indicates tone presentation. Bottom, raster plot, colour-coded as oddball (green) and standard (black). **f**, The mean \pm s.e.m. response across all units ($n = 150$ units, 2 patients) to oddball (green) versus standard (black) tones. The red line indicates significantly different periods. **g**, The average ERP. $n = 2$ patients, 10 channels each. **h**, The average gamma band amplitude. **i**, β -Coefficients from a linear regression run

per unit, with modulation a function of tone identity (purple), oddball identity (green) and an interaction term (yellow). Statistical analysis was performed using t -tests for difference in absolute amplitude. $n = 150$ units, 2 patients for subsequent analyses. The box limits show the interquartile range (IQR); the centre dot shows the median and the whiskers show $1.5 \times$ IQR past the box. The lines are the kernel density estimate. **j**, The decoding accuracy for tone (purple) and oddball identity (green) for p5 and p6 across units (left) and ERP (right) after tone presentation. The whiskers here are the extrema within 150% of the IQR past the box, the plus symbols represent outliers. Statistical analysis was performed using a one-sided t -test. **k**, The SVM decoding accuracy for tone identity (purple) and oddball identity (green) across all channels within each frequency band. Statistical analysis was performed using a one-sided t -test. No adjustments made for multiple comparisons. $*P < 0.05$, $***P < 0.001$.

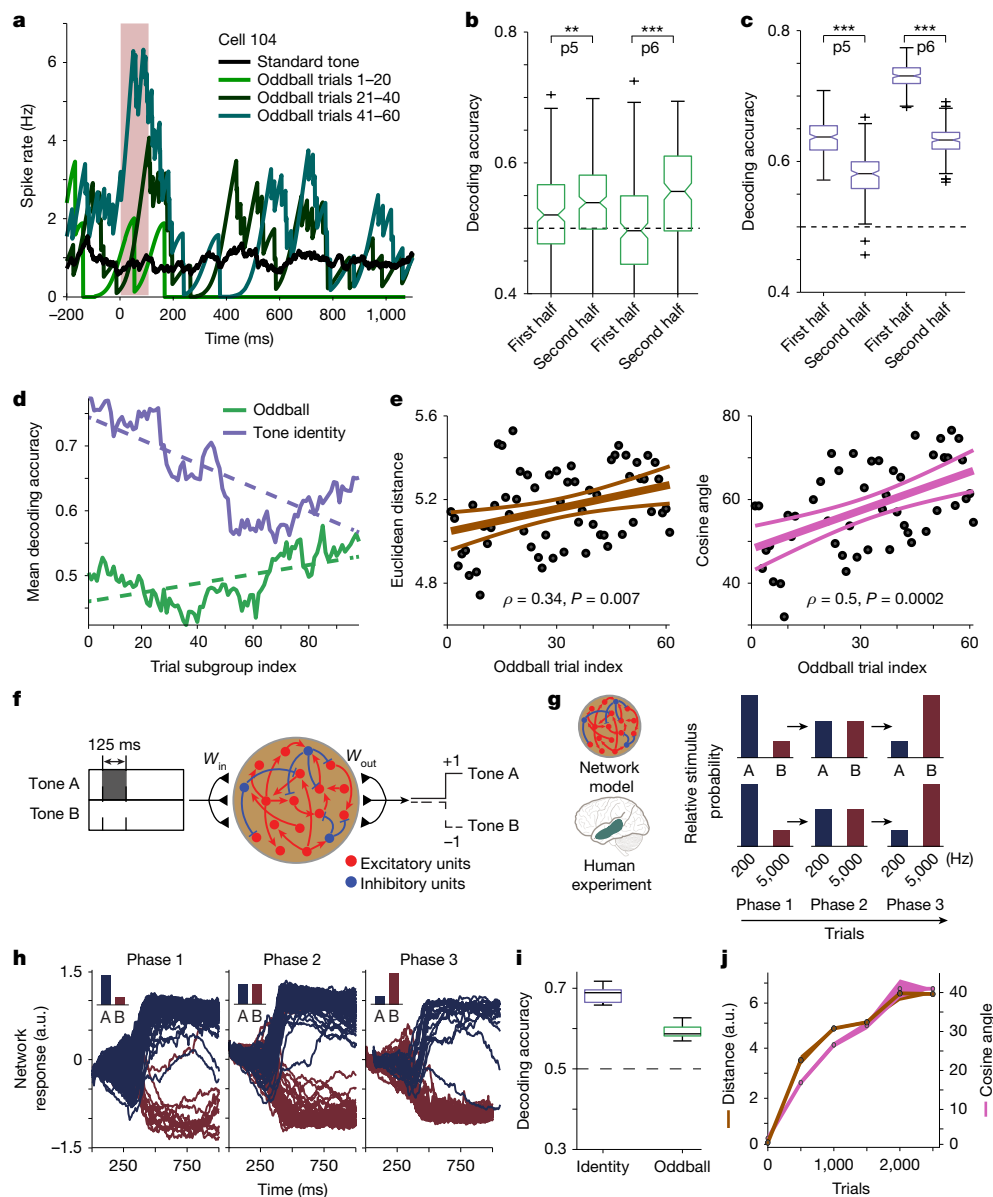


Fig. 3 | Evolution of oddball representation across the neuronal population in experimental data and an RNN model. a, Tone responses as a function of oddball identity and index in an example unit. The red bar indicates tone presentation. **b**, The accuracy of neuronal population oddball decoding for patients p5 and p6 ($n = 43$ oddball-responsive units), for the first half of each block (left) or second half of each block (right), combined across blocks. The boxes represent the 25th–75th percentiles, notched at the median. The whiskers are the extrema within 150% of the IQR past the box, plus symbols represent outliers. Statistical analysis was performed using two-sided paired t -tests; $**P = 0.004$. **c**, Similar to **b** but for tone identity. **d**, Decoding accuracy for tone (purple) and oddball (green) identity as a function of trial position. Each point represents SVM accuracy within a set of 50 trials starting at the index location. The dashed lines are linear fits. **e**, The Euclidean distance (left)

and cosine angle (right) between standard and oddball neuronal population response vectors, computed for each oddball trial. The lines are linear fits with the 95% confidence intervals. **f**, Schematic of the RNN model trained to differentiate between two different tone frequencies, indicated as tone A and tone B. **g**, The training paradigm for the RNN as compared with the human experiment. **h**, Network response across trials for a single example RNN model unit for the three training phases. a.u., arbitrary units. **i**, The decoding accuracy of RNN for tone identity (purple) and oddball identity (green) ($n = 100$ generated trials; Methods). The boxes represent the 25th–75th percentiles, notched at the median; the whiskers are the extrema. **j**, Evolution of Euclidean distance (brown) and cosine angle (pink) between oddball trials and the average standard trial across the RNN population. The shading represents the s.e.m. across ten runs. No adjustments were made for multiple comparisons.

These results indicate that the hippocampus does not simply improve encoding using gain modulation¹⁸; instead, oddball responses reflect a rotation of the neural population vector in a high-dimensional space, meaning that neural plasticity alters the shape of the neural response manifold¹⁹. Thus, complex reshaping of responses can occur even under general anaesthesia. Further analyses of the information encoded in LFP, including band-limited analysis and aperiodic slope is provided in the Supplementary Notes and Extended Data Figs. 1–3. Additional analyses relating to cell-type encodings can be found in Extended Data Fig. 4.

To gain further mechanistic insights at the level of individual units, we turned to a continuous-rate recurrent neural network (RNN) trained to perform a signal-detection task similar to the task used for the human Neuropixels data²⁰ (Fig. 3f). The network model underwent three stages of training, simulating the different contexts used in the experimental data, with the prevalence of specific tones varied at each stage (Fig. 3g,h and Methods). To simulate the two auditory tones, the model received Gaussian white noise across two input channels, each corresponding to one of the two tones (tone A and tone B). On each trial, a transient

+1.0 bias was added to one of the two channels during a brief stimulus window, indicating the presence of the corresponding tone.

Tone A was presented to the network in 80% of the trials, followed by a washout period and then a stage with the probabilities reversed relative to the first. By the end of training (range of 1,400 to 2,600 trials), the model was able to differentiate between tone identities (Fig. 3h). Notably, despite being explicitly trained only on tone identity discrimination, the model was able to perform not only identity discrimination (tone identity, $P < 0.005$, signed Wilcoxon test versus shuffled data) but also context classification (oddball identity, $P < 0.005$, signed Wilcoxon test versus shuffled data; Fig. 3i) through linear SVM decoding (Methods). The model also recreated the pattern observed in the Euclidean and vector angle distance between standard and oddball representations (Fig. 3j), suggesting that the divergence of representations can be due to local computations rather than inherited from other networks. Note that, although the standard tone occurs in 80% of trials, the network was trained to make a binary choice between two alternatives. Thus, the theoretical chance performance remains 50%.

Rate-based RNNs provide a tractable framework for investigating how excitatory–inhibitory (E–I) interactions give rise to these mechanisms. To this end, we used an E–I RNN model to reflect the known composition of cortical circuits, which are predominantly excitatory (around 80%) with a smaller proportion of inhibitory neurons (around 20%). To more fully leverage the E–I structure of the model, we conducted systematic lesioning analyses in which we selectively lesioned each of the four recurrent connection subtypes (E to I, E to E, I to E, I to I) by setting the corresponding weights to zero. We then recomputed SVM decoding performance for both tone identity and oddball context.

Systematic lesioning of each of the four recurrent synaptic connection types (E to E, E to I, I to E, I to I) revealed that inhibitory connections are essential for encoding both tone and oddball categories (Extended Data Fig. 5), as lesioning I-to-E and I-to-I connections led to the most pronounced decrease in the decoding accuracy for both tone identity and oddball context. These findings indicate that inhibitory feedback, both directly onto excitatory neurons and within inhibitory populations, has a critical role in shaping population-level representations. Inhibitory connections were important for encoding both tone identity and oddball context.

Language in the unconscious hippocampus

We next tested whether the unconscious hippocampus could perform even higher order functions associated with parsing semantic and syntactic features of natural speech. In four participants (p6, p8, p9 and p11), we recorded neural activity while playing 10–20 min of podcast episodes²¹ (Methods). We aligned neural activity to word onset and offset ($n = 3,024$ words for p6; $n = 1,565$ words for p8 and p9; $n = 962$ words for p11), and computed word-evoked neural responses (example unit, average response to all words presented; Fig. 4a).

Given the oddball effects described above, we first hypothesized that the brain would respond differentially based on word lexical frequency, which we defined using a standard database²². We found a statistically significant correlation between the firing rate and word frequency in all four patients individually (Fig. 4b and Extended Data Fig. 6). Specifically, we found a significant positive correlation in the single-unit activity (mean $r = 0.48 \pm 0.06$, Spearman's correlation, $\alpha < 0.05$) and across patients ($P < 0.0001$, GLME) and a modest but significantly negative correlation in all six bands of the LFP (range -0.08 in the beta band to -0.02 in the delta band, $P < 0.001$ for all). Notably, lexical duration is also significantly encoded ($P < 0.0001$, GLME). To address possible confounding factors between word duration and frequency, we reran the unit analysis with subsets of words within a limited duration range, that is, 0–200 ms, 200–400 ms and 400–600 ms, and consistently

observed a positive correlation ($P < 0.001$ for combined units in each patient separately). Moreover, a linear model that incorporated both logarithmic word duration and logarithmic word frequency still found significance in word frequency as a predictor of firing rate ($P < 0.001$, in each patient separately, t -test on coefficients). This correlation could not be solely explained by difficulty in lexical access, as there was also a consistent relationship of the neural responses with the relative surprise of each word (see below).

These results suggest that the unconscious hippocampus has access to the semantic information conveyed by each word. To explicitly test this possibility, we regressed the firing rates of each neuron against the semantic embeddings of each word that demonstrated a response^{21,23} (Methods). In semantic embedding space, similar words (such as 'dog' and 'cat') are closer (Euclidean distance, $d = 1.8$) than dissimilar words (for example, 'dog' and 'pen', $d = 2.5$). Using tenfold cross-validation, we found that the root-mean-squared error (r.m.s.e.) of a linear model outperformed shuffled data in all units ($\alpha = 0.05$, one tailed t -test on real versus shuffled r.m.s.e.; Fig. 4c), with an average correlation between true and predicted firing rates of 0.397 ± 0.007 ($n = 368$ units).

Overall, these results parallel those obtained from a separate cohort of awake patients who performed a similar task in the EMU with single units recorded on microwire electrodes²¹. Specifically, in awake patients, the mean correlation was slightly lower, at 0.226 ± 0.009 (356 units across 10 patients). However, given that conversational English has many words that are repeated, these results could theoretically be confounded by the fact that cells had consistent responses to words, perhaps even matching acoustic features. To show that units generalize across word embeddings, we re-ran the analysis using only unique words ($n = 743, 571, 571$ and 329 words for the four patients, respectively). We found a significant result in 75.4% of the recorded units (251 out of 333 units with at least 50 words that had a non-zero response), with an average correlation of $r = 0.207 \pm 0.05$; Fig. 4d). These numbers were again comparable to the numbers observed in the awake cohort (73.2% of units; 246 out of 336; mean $r = 0.134 \pm 0.005$). In other words, it is possible to predict the firing rate of units to a given word based on responses to other words by leveraging their similarities in semantic space²⁴, demonstrating that the unconscious hippocampus has access to abstract semantic relationships between words.

We then analysed the representation of word features. We placed each word into one of twelve semantic categories²¹ (Fig. 4e). Nearly all units (85.6%, $n = 321$ out of 375) showed some form of semantic category selectivity ($\alpha = 0.05$, Kruskal–Wallis test for any difference between semantic categories; Fig. 4f). This number is similar to the corresponding number in awake patients (76.1%, $n = 271$ out of 356). Units were selective for multiple semantic categories, consistent with our previously reported findings in awake patients (rank-sum test, corrected for multiple comparisons, $\alpha < 0.05$)²¹. Specifically, 239 out of 375 (63.7%) units discriminated at least 2 out of 12 categories; 54 out of 375 (14.4%) discriminated at least 4 (Fig. 4g), median of 3 categories per neuron. The corresponding numbers in awake patients were 232 out of 356 and 139 out of 356, respectively. We next found that 298 out of 375 units carried information about part of speech²⁵ ($\alpha = 0.05$; Kruskal–Wallis test; Fig. 4h), with nearly identical numbers for awake patients: 298 out of 356. Again, there was broad representation of different categories (Fig. 4i). Most units (80.0%) distinguished nouns from non-nouns, but none distinguished verbs from non-verbs; these numbers were comparable in awake patients (79.5% and 0%). Overall, the median number of part of speech categories represented was 4 out of 11 (Fig. 4j), with 259 out of 375 (69.1%) units discriminating at least two categories and 106 out of 375 (28.3%) discriminating at least four. These numbers were similar for awake patients: 251 out of 356 (70.5%) units discriminated at least 2 categories; 141 out of 356 (39.6%) units discriminated at least 4 categories. Notably, we found a modest, but significant correlation between the number of semantic

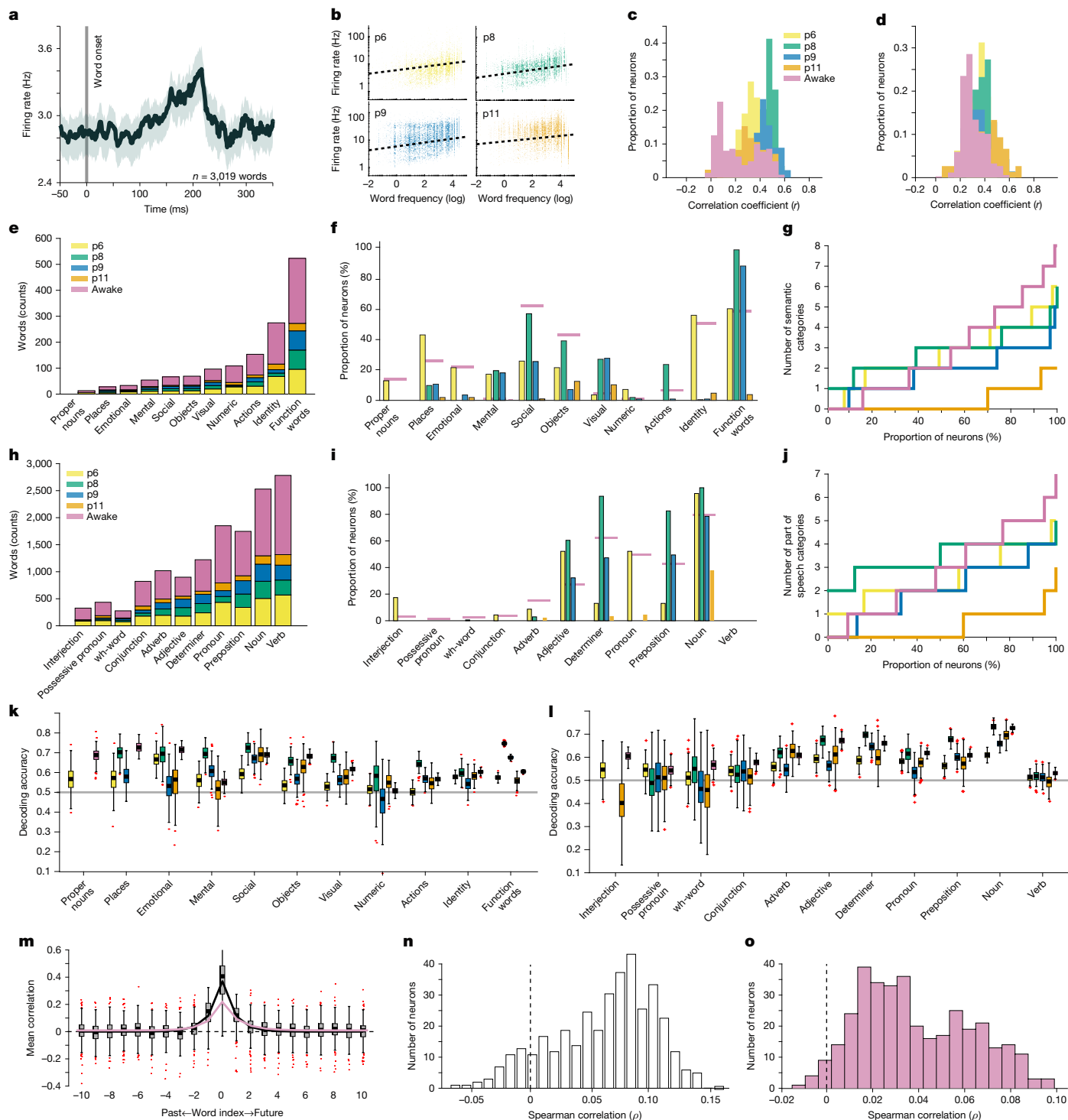


Fig. 4 | Language responses in the anaesthetized human hippocampus.

a, The response (spike rate, Hz) of an example unit, aligned to word onset (vertical line). Data are mean \pm s.e.m. **b**, Neuronal responses as a function of word frequency shown for each neuron. Each datapoint corresponds to a neuron and word, colour-coded according to the patient. The dashed line shows the linear fit for that patient. **c, d**, The distribution of the Pearson's correlation coefficient for predicted versus actual firing rates for all (c) and unique (d) words, per patient. All correlations are greater than zero, indicating that the models are better than chance at fitting neural responses. **e**, The distribution of words within semantic categories sorted by frequency within the stimulus set, per patient. **f**, The percentage of units selective for each semantic category, compared with all other semantic categories, per patient. The horizontal pink lines represent values from awake data. **g**, The number

of categories decoded by individual units, per patient. Neurons generally participate in coding of multiple categories. **h-j**, Similar to **e-g**, respectively, but for part of speech categories. **k, l**, The decoding accuracy of a SVM across units for semantic (k) and part of speech (l) categories, shown per patient. The grey lines represent chance; the boxes represent the 25th–75th percentiles, notched at the median. The whiskers are the extrema within 150% of the IQR past the box. The plus symbols indicate outliers. $n = 4$ patients, 375 units. **m**, Correlation coefficients for predicted versus actual firing rates as a function of word index. Index = 0, current word. Negative indices indicate past words; positive indices indicate future words. Black, anaesthetized ($n = 4$ patients, 375 units); pink, awake ($n = 10$ patients, 356 units). **n, o**, The correlation coefficient (Spearman) for each neuron between surprisal and firing rate in anaesthetized (n) and awake (o) patients.

categories and the number of part of speech categories represented across neurons ($r = 0.34$, $P < 0.001$ Spearman's correlation; awake: $r = 0.24$; $P < 0.01$), suggesting that language-responsive neurons can represent multiple features. Word frequency encoding effects are robust using $P < 0.01$ (98.1%; awake patients: 86.8%) and at $P < 0.001$ (97.1%; awake patients: 79.8%). Semantic encoding proportions are also robust using $P < 0.01$ (77.9% and 68.5%) and $P < 0.001$ (62.1% and 52.0%). Likewise, part of speech was $P < 0.01$ (71.5% and 62.1%) and $P < 0.001$ (74.4% and 62.6%).

We next examined hippocampal decoding ability on a word-by-word basis. We used a SVM to compare each category against all others. We found that all categories in both semantics and part of speech were decodable at the level of individual words ($P < 0.001$ versus shuffled data, Fig. 4k,l; note that chance is 0.5 because we subsampled the majority class such that positive and negative data had equal frequency). Semantic categories had a higher average classification accuracy ($60.5 \pm 4.0\%$) than part of speech categories ($56.5 \pm 5.3\%$, $P = 0.03$, t -test). Thus, both semantic and syntactic information is decodable in real time within the anaesthetized hippocampus.

We next examined whether the anaesthetized hippocampus could represent recent or upcoming words²⁶. We reran our linear regressions predicting previous and upcoming words (using held-out data to prevent overfitting; Methods). We found that neural responses corresponded not only to semantic features of previous words (Fig. 4m; negative indices), which could be due to short term memory²⁷, but also to the semantics of future words²⁸ (Fig. 4m; positive indices). Future words were decoded nearly as well as past words (across all four patients: $\beta_0 = 0.370$, $\tau_{\text{future}} = 0.840$; $\tau_{\text{past}} = 0.868$). These numbers are comparable to those for awake patients ($\beta_0 = 0.227$; $\tau_{\text{future}} = 1.081$; $\tau_{\text{past}} = 0.895$); indeed, they do not differ significantly ($P > 0.05$ in all cases in the range +1 to +5). Thus, not only is recent speech actively maintained, but encoding of words is contextualized such that we can decode future words from these encodings; this type of contextualization is crucial to speech comprehension²⁹. However, these results do not necessarily imply active prediction beyond contextualization³⁰. Moreover, firing rates of many units were modulated by surprisal, a metric that quantifies the probability of each word as a function of prior words³¹ ($r = 0.06 \pm 0.0023$, 246 out of 375 units significant; Fig. 4n,o). Similar results were observed in awake patients ($r = 0.0386 \pm 0.0013$, 245 out of 356 neurons significant). Surprisal effects are also robust at $P < 0.01$ and $P < 0.001$ in anaesthetized (56.8% and 40.3%, respectively) and in awake patients (56.2% and 43.0%, respectively).

Further analyses of the LFP are shown in Extended Data Figs. 7–9. In brief, we found that semantics were encoded in all six bands (Extended Data Fig. 7), although less faithfully than in unit data (Fig. 4c). We also found encoding of semantic category in all bands (Extended Data Fig. 8a) and, again, less strongly than with the single units (Fig. 4f,g). Likewise, we found significant encoding of part of speech in all frequency bands (Extended Data Fig. 8b), although less strongly than semantic category (Fig. 4i,j). An SVM approach was able to classify part of speech and semantic category (Extended Data Fig. 9) about as well as the single units in all bands (Fig. 4k,l). Aperiodic slope outperformed individual spectral features in predicting semantic embeddings (mean $r = 0.32 \pm 0.17$, rank-sum test, $P < 0.001$ for all bands) and robustly encoded multiple semantic and POS categories (62.5% and 25.0% encoded at least 8 categories, respectively).

We also found robust phase-amplitude coupling results for the language data. Theta-low gamma phase-amplitude coupling outperformed spectral features in predicting semantic embeddings (mean $r = 0.170 \pm 0.13$, rank-sum test, $P < 0.001$ for all bands) and similarly encoded multiple semantic and POS categories (74.5% and 73.0% encoded at least 8 categories, respectively). Phase-amplitude coupling using the high gamma band (70–150 Hz) showed comparable semantic embedding prediction (mean $r = 0.166 \pm 0.13$) and improved

encoding of semantic and POS categories (99.3% and 98.7% encoded at least 8 categories, respectively). However, spike-frequency coupling in the theta band did not significantly encode semantic embeddings (mean $r = 0.002 \pm 0.004$) or categories (7% and 2% of units encoded one semantic and POS category, respectively; none encoded 2 or more). Together, these results attest to the robust recoverability of linguistic information in the anaesthetized brain.

Discussion

Here we identified neural signatures of plasticity and semantic processing in the anaesthetized human hippocampus. Our findings do not have obvious explanations based solely on low-level sensory responses. For example, the long and slow increase in oddball detection over the course of 10 min is unlikely to reflect adaptation or repetition suppression, which both take place on shorter timescales. Likewise, the representation of semantic features in speech listening requires specific processing of acoustic information. We therefore show that, within anaesthetic-induced unconsciousness, some high-level process of sensory integration is preserved, suggesting that it is consolidation that is compromised. These results provide a potential explanation for previous reports of post-anaesthesia implicit recall, which would depend on sensory processing and plasticity processes³².

Broadly speaking, these results confirm ideas previously developed in animal and human studies showing preserved neural responses to sensory stimuli, including oddball stimuli, during anaesthesia³³, and extend them to include (1) change over the timescale of several minutes, a timescale usually associated with wakeful learning; (2) linkage to a plausible biocomputational model that avoids the types of executive control that are presumably diminished during anaesthesia; and (3) availability of language information beyond the level of auditory processing. The present results complement our past studies, which were performed using awake patients in the epilepsy monitoring unit (EMU), on hippocampus neuron-level representation of lexical semantics²¹ and semantic contextualization³⁴. These results suggest that awake-like semantic responses, and at least some contextualization, can occur in the absence of conscious awareness. That in turn raises the question of how far linguistic processing can go in the absence of awareness¹⁰.

Some aspects of the local field potential correlate with single-unit activity, although this relationship may fluctuate³⁵. Much evidence suggests that the gamma band in particular may align closely with single units³⁶. We find that the gamma band is generally aligned with neural activity, although not perfectly so, and that other bands are sometimes also aligned. However, failure to achieve significant effects may reflect an insufficiency of data; it is therefore difficult to draw firm negative conclusions from observed differences between unit and LFP responses. In any case, our results do indicate that LFP activity, especially in the gamma band, can be a proxy for single units, although it may have other functional roles as well.

While the hippocampus is not a well-known part of the core language network, it has a well-established role in language³⁷, including in recent single-unit recording studies^{21,38}. In particular, its known functions echo semantic contextualization: (1) it is associated with mapping functions related to temporal position encoding³⁹; (2) it integrates information across multiple modalities and uses that information to contextualize representations in a variety of domains⁴⁰; (3) and it is closely associated with the processes of prediction that drive contextualization⁴¹; and, finally (4) it has a prominent role in memory⁴². One possible reason for the absence of the hippocampus in classical language models is negative evidence—it is difficult to image using functional magnetic resonance imaging and lesions isolated to the hippocampus are rare.

Our study has several limitations. First, anaesthesia has an uncertain relationship with waking life⁴³. Moreover, it remains unclear whether

our findings will apply to other non-conscious states, such as sleep and coma. Indeed, our results may not generalize to anaesthetics besides propofol. Furthermore, we did not have enough patients to test for lateralization/dominance. Another limitation is that processes that we describe may not be unique to the hippocampus. Indeed, the RNN relies critically on the formatting of its input into binary categories, which must be inherited from other regions. Instead, the goal is to understand how oddball-dependent responses can arise from such input. Furthermore, we cannot definitively conclude that plastic changes in tone identity decoding during the oddball task are a compensatory property of the oddball response, rather than an independent plastic effect such as stimulus-specific adaptation. Finally, our choice to use temporally unpredictable tones may have weakened our sensitivity to tone-related effects as temporally predictable oddballs may maximally drive neural responding⁴⁴.

These results highlight the robust coding of cognitive variables in the hippocampus in the absence of consciousness. Such task-correlated patterns of neural activity are typically thought of as neural correlates of their corresponding cognitive processes, so our observations raise the possibility that those processes may occur in the absence of consciousness. That idea, in turn, suggests that consciousness may have some association with those processes but is not essential. Instead, those processes may occur in a subconscious manner, as for example, we may monitor subconsciously others' conversations at a cocktail party⁴⁴. A good deal of past research has emphasized the central positioning of the hippocampus within the anatomical hierarchy of brain areas; indeed, it may be among the regions most distant from both input and output ends of brain processing⁴⁵. It is therefore generally assumed that the hippocampus will have the most attenuated inputs in the absence of consciousness; our results invite a reconsideration of that idea.

These results raise the question about what features differentiate conscious from non-conscious processing. Our results, at least, suggest that the key ingredient does not reside in the activity of single units or LFP in the hippocampus. Existing prominent theories consistent with our results include the idea that consciousness involves cross-regional coordination^{8,46}, global propagation of local signals^{47,48} or recurrent processing⁴⁹. Another possibility is that consciousness is not simply a result of neural activity at a given time but is instead constructed through repeated revisions of current experiences, and that it is this revision process that is impaired⁵⁰.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-026-10448-0>.

- van Gaal, S., De Lange, F. P. & Cohen, M. X. The role of consciousness in cognitive control and decision making. *Front. Hum. Neurosci.* **6**, 121 (2012).
- Dehaene, S., Lau, H. & Kouider, S. What is consciousness, and could machines have it? *Science* **358**, 486–492 (2017).
- Sikkens, T., Bosman, C. A. & Olcese, U. The role of top-down modulation in shaping sensory processing across brain states: implications for consciousness. *Front. Syst. Neurosci.* **13**, 31 (2019).
- Alkire, M. T., Hudetz, A. G. & Tononi, G. Consciousness and Anesthesia. *Science* **322**, 876–880 (2008).
- Brown, E. N., Lydic, R. & Schiff, N. D. General anesthesia, sleep, and coma. *N. Engl. J. Med.* **363**, 2638–2650 (2010).
- Coughlin, B. et al. Modified Neuropixels probes for recording human neurophysiology in the operating room. *Nat. Protoc.* **18**, 2927–2953 (2023).
- Hickok, G. & Poeppel, D. The cortical organization of speech processing. *Nat. Rev. Neurosci.* **8**, 393–402 (2007).
- Dehaene, S. & Changeux, J.-P. Experimental and theoretical approaches to conscious processing. *Neuron* **70**, 200–227 (2011).
- Tononi, G., Boly, M., Massimini, M. & Koch, C. Integrated information theory: from consciousness to its physical substrate. *Nat. Rev. Neurosci.* **17**, 450–461 (2016).
- Kouider, S. & Dehaene, S. Levels of processing during non-conscious perception: a critical review of visual masking. *Philos. Trans. R. Soc. Lond. B* **362**, 857–875 (2007).
- Sklar, A. Y. et al. Reading and doing arithmetic nonconsciously. *Proc. Natl Acad. Sci. USA* **109**, 19614–19619 (2012).
- Hassin, R. R. Yes it can: on the functional abilities of the human unconscious. *Perspect. Psychol. Sci. J. Assoc. Psychol. Sci.* **8**, 195–207 (2013).
- Chung, J. E. et al. High-density single-unit human cortical recordings using the Neuropixels probe. *Neuron* **110**, 2409–2421 (2022).
- Tzovara, A. et al. Predictable and unpredictable deviance detection in the human hippocampus and amygdala. *Cereb. Cortex* **34**, bhad532 (2024).
- Billig, A. J., Lad, M., Sedley, W. & Griffiths, T. D. The hearing hippocampus. *Prog. Neurobiol.* **218**, 102326 (2022).
- Kelemen, E. & Fenton, A. A. Dynamic grouping of hippocampal neural activity during cognitive control of two spatial frames. *PLoS Biol.* **8**, e1000403 (2010).
- Fontanini, A. & Katz, D. B. Behavioral states, network states, and sensory response variability. *J. Neurophysiol.* **100**, 1160–1168 (2008).
- Treue, S. & Maunsell, J. H. Effects of attention on the processing of motion in macaque middle temporal and medial superior temporal visual cortical areas. *J. Neurosci.* **19**, 7591–7602 (1999).
- Ebitz, R. B. & Hayden, B. Y. The population doctrine in cognitive neuroscience. *Neuron* **109**, 3055–3068 (2021).
- Rungtameeetaeuewana, N., Kim, R., Chotitub, T. & Sejnowski, T. J. Random noise promotes slow heterogeneous synaptic dynamics important for robust working memory computation. *Proc. Natl Acad. Sci. USA* **122**, e2316745122 (2025).
- Franch, M. et al. A vectorial code for semantics in human hippocampus. Preprint at *bioRxiv* <https://doi.org/10.1101/2025.02.21.639601> (2025).
- Brysaert, M. & New, B. Moving beyond Kučera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behav. Res. Methods* **41**, 977–990 (2009).
- Jamali, M. et al. Semantic encoding during language comprehension at single-cell resolution. *Nature* **631**, 610–616 (2024).
- Khanna, A. R. et al. Single-neuronal elements of speech production in humans. *Nature* **626**, 603–610 (2024).
- Manning, C. D. et al. The Stanford CoreNLP Natural Language Processing Toolkit. In *Proc. 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations* (eds Bontcheva, K. & Zhu, J.) 55–60 (Association for Computational Linguistics, 2014).
- Heilbron, M., Armeni, K., Schoffelen, J.-M., Hagoort, P. & de Lange, F. P. A hierarchy of linguistic predictions during natural language comprehension. *Proc. Natl Acad. Sci. USA* **119**, e2201968119 (2022).
- Jang, A. I., Wittig, J. H. Jr, Inati, S. K. & Zaghoul, K. A. Human cortical neurons in the anterior temporal lobe reinstate spiking activity during verbal memory retrieval. *Curr. Biol.* **27**, 1700–1705 (2017).
- Pickering, M. J. & Gambi, C. Predicting while comprehending language: a theory and review. *Psychol. Bull.* **144**, 1002–1044 (2018).
- Ryskin, R. & Nieuwland, M. S. Prediction during language comprehension: what is next? *Trends Cogn. Sci.* **27**, 1032–1052 (2023).
- Schönmann, I., Szewczyk, J., de Lange, F. P. & Heilbron, M. Stimulus dependencies—rather than next-word prediction—can explain pre-onset brain encoding during natural listening. *eLife* **14**, RP106543 (2025).
- Weissbart, H., Kandylaki, K. D. & Reichenbach, T. Cortical tracking of surprisal during continuous speech comprehension. *J. Cogn. Neurosci.* **32**, 155–166 (2020).
- Linassi, F. et al. Implicit memory and anesthesia: a systematic review and meta-analysis. *Life* **11**, 850 (2021).
- Nourski, K. V. et al. Auditory predictive coding across awareness states under anesthesia: an intracranial electrophysiology study. *J. Neurosci.* **38**, 8441–8452 (2018).
- Katlowitz, K. A. et al. Attention is all you need (in the brain): semantic contextualization in human hippocampus. Preprint at *bioRxiv* <https://doi.org/10.1101/2025.06.23.661103> (2025).
- Ojemann, G. A., Ramsey, N. F. & Ojemann, J. Relation between functional magnetic resonance imaging (fMRI) and single neuron, local field potential (LFP) and electrocorticography (ECoG) activity in human cortex. *Front. Hum. Neurosci.* **7**, 34 (2013).
- Henrie, J. A. & Shapley, R. LFP power spectra in V1 cortex: the graded effect of stimulus contrast. *J. Neurophysiol.* **94**, 479–490 (2005).
- Duff, M. C. & Brown-Schmidt, S. The hippocampus and the flexible use and processing of language. *Front. Hum. Neurosci.* **6**, 69 (2012).
- Dijksterhuis, D. E. et al. Pronouns reactivate conceptual representations in human hippocampal neurons. *Science* **385**, 1478–1484 (2024).
- Naya, Y. & Suzuki, W. A. Integrating what and when across the primate medial temporal lobe. *Science* **333**, 773–776 (2011).
- Maren, S., Phan, K. L. & Liberzon, I. The contextual brain: implications for fear conditioning, extinction and psychopathology. *Nat. Rev. Neurosci.* **14**, 417–428 (2013).
- Aitken, F. & Kok, P. Hippocampal representations switch from errors to predictions during acquisition of predictive associations. *Nat. Commun.* **13**, 3294 (2022).
- Binder, J. R. & Desai, R. H. The neurobiology of semantic memory. *Trends Cogn. Sci.* **15**, 527–536 (2011).
- Franks, N. P. & Lieb, W. R. Mechanisms of general anesthesia. *Env. Health Perspect.* **87**, 199–205 (1990).
- Bronkhorst, A. W. The cocktail-party problem revisited: early processing and selection of multi-talker speech. *Atten. Percept. Psychophys.* **77**, 1465–1487 (2015).
- Margulies, D. S. et al. Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proc. Natl Acad. Sci. USA* **113**, 12574–12579 (2016).
- Mashour, G. A., Roelfsema, P., Changeux, J.-P. & Dehaene, S. Conscious processing and the global neuronal workspace hypothesis. *Neuron* **105**, 776–798 (2020).

47. Baars, B. J. Global workspace theory of consciousness: toward a cognitive neuroscience of human experience. *Prog. Brain Res.* **150**, 45–53 (2005).
48. Dehaene, S., Kerszberg, M. & Changeux, J. P. A neuronal model of a global workspace in effortful cognitive tasks. *Proc. Natl Acad. Sci. USA* **95**, 14529–14534 (1998).
49. Lamme, V. A. F. Towards a true neural stance on consciousness. *Trends Cogn. Sci.* **10**, 494–501 (2006).
50. Dennett, D. C. Real patterns. *J. Philos.* **88**, 27–51 (1991).
51. Magnotti, J. F., Wang, Z. & Beauchamp, M. S. RAVE: Comprehensive open-source software for reproducible analysis and visualization of intracranial EEG data. *NeuroImage* **223**, 117341 (2020).

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



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026

Methods

Patient recruitment

Experiments were conducted according to protocol guidelines approved by the Institutional Review Board for Baylor College of Medicine and Affiliated Hospitals (H-50885 for the Neuropixels recordings and H-18112 for the EMU recordings). All of the recruited patients for the Neuropixels recordings were diagnosed with drug-resistant temporal lobe epilepsy and were scheduled to undergo an anteromesial temporal lobectomy for seizure control. All of the patients provided written informed consent to participate in the study and were aware that participation was voluntary and would not affect their clinical course. Included patients' age ranged from 25–54 years old (average, 39.6 ± 11.8), with three female and four male patients. Four resections were on the left side, and three were on the right. In one individual (p3), recordings were performed in the middle temporal lobe before resection. None of the patients reported explicit memory of intraoperative events after the case when discussed in the post-operative care unit or while recovering in the hospital the next day.

Note that we include for comparison purposes a cohort of awake patients listening to podcast stimuli. These patients were recruited from patients undergoing invasive recordings in the EMU at Baylor St Luke's Hospital. Details on methods for this group of patients were reported previously^{21,34,52–54}.

Neuropixels data acquisition set-up and intraoperative recordings

Neuropixels 1.0-S probes (IMEC) with 384 recording channels (total recording contacts = 960, usable recording contacts = 384) were used for recordings (dimensions: 70 μm width, 100 μm thickness, 10 mm length). The Neuropixels probe, consisting of both the recording shank and the headstage, were individually sterilized with ethylene oxide (Bioseal)⁶. Our intraoperative data acquisition system included a custom-built rig including a PXI chassis affixed with an IMEC/Neuropixels PXIe Acquisition module (PXIe-1071) and National Instruments DAQ (PXI-6133) for acquiring neuronal signals and any other task-relevant analogue/digital signals respectively. Our recording rig was certified by the Biomedical Engineering at Baylor St Luke's Medical Center, where the intraoperative recording experiments were conducted. A high-performance computer (10-core processor) was used for neural data acquisition using open-source software such as SpikeGLX 3.0 and OpenEphys v.0.6x for data acquisition (the action potential (AP) band was band-pass filtered from 0.3 kHz to 10 kHz and acquired at 30 kHz sampling rate; the LFP band was band-pass filtered from 0.5 Hz to 500 Hz and acquired at a 2,500 Hz sampling rate). We used a short-map probe channel configuration for recording, selecting the 384 contacts located along the bottom third of the recording shank.

Audio was played through a separate computer using pregenerated .wav files and captured at 30 kHz or 1,000 kHz on the NIDAQ through a coaxial cable splitter that sent the same signal to speakers adjacent to the patient. MATLAB (MathWorks) in conjunction with a LabJack (LabJack U6) was used to generate a continuous TTL pulse of which the width was modulated by the current timestamp and recorded on both the neural and audio datafiles. Online synchronization of the AP and LFP files was performed by the OpenEphys recording software. Offline synchronization of the neural and audio data was performed by calculating a scale and offset factor via a linear regression between the time stamps of the reconstructed TTL pulses and confirmed with visual inspection of the aligned traces.

Acute intraoperative recordings were conducted in brain tissue designated for resection based on purely clinical considerations. The probe was positioned using a ROSA ONE Brain (Zimmer Biomet) robotic arm and lowered into the brain 5–6 mm from the ependymal surface using an AlphaOmega (Alpha Omega Engineering). The penetration was monitored through online visualization of the neuronal data and

through direct visualization with the operating microscope (Kinevo 900). Reference and ground signals on the Neuropixels probe were acquired by connecting to sterile needles placed in the scalp (separate needles inserted at distinct scalp locations for ground and reference respectively).

For all patients ($n = 7$), we conducted neuronal recordings under general anaesthesia for at most 30 min as per the experimental protocol. All of the patients were under total intravenous anaesthesia, with propofol as the main anaesthetic for each experimental protocol (Extended Data Table 1). Inhaled anaesthetics were only used for induction and stopped at least 1 h before recordings. The anaesthesiologist titrated the anaesthetic drug infusion rates so that the BIS monitor (Medtronic) value was between 45 and 60 for the duration of the surgical case⁵⁵. Notably, BIS values range between 0 (completely comatose) and 100 (fully awake), with standard intraoperative values between 40 and 60. Specific anaesthesia depth was flat across the brief time of the experiment. First, recordings took place several hours after anaesthesia induction and several hours before the end of the procedure, so patients were well into the stable portion of the surgery. Second, the anaesthesiologist was maintaining active monitoring and stably controlled anaesthesia levels.

For patients p4, p5 and p6, we recorded neuronal activity during passive auditory stimuli presentation. For p4, a sequence of auditory stimuli (pure tones; $f_1 = 1$ kHz, $f_2 = 3$ kHz) was presented with an 80–20 probability distribution, with the less frequent auditory stimulus serving as an auditory oddball stimulus ($n = 300$ trials). For p5 and p6 we counterbalanced the tones. A sequence of auditory stimuli (pure tones; $f_1 = 200$ Hz, $f_2 = 5$ kHz) were presented with an 80–20 probability distribution, while switching the tone frequency designated as the auditory oddball stimulus halfway through (first half, $n = 150$ trials, f_2 is auditory oddball; second half, $n = 150$ trials, f_1 is auditory oddball). We interleaved a washout period (30 trials) using the same auditory stimuli but presented at a 50–50 probability distribution in between the two counterbalanced tasks. The auditory pure tone stimuli were presented for a 100 ms duration, and the intertrial interval for the auditory oddball task was randomly drawn from between 1 and 3 s.

Sound stimuli for the auditory oddball task consisted of high- and low-pitched tones. The low-pitched tone was 100 ms duration and 200 Hz, approximating a square wave. The high-pitched tone was 100 ms duration and 5 kHz frequency, also approximating a square wave. These stimuli were constructed to have distinct perceived pitch and salient onset structure. Stimulus waveforms were matched in amplitude. Sounds were delivered in stereo, using a sound delivery system that was calibrated in the testing suite (B&K type 4939-A-011 calibration mic and NEXUS 4939-A-011 conditioning amplifier). Both speakers had a relatively flat frequency response (± 5 dB) across the used frequency range (200–6,000 Hz) and no high- or low-frequency roll-off.

For patients p6, p8, p9 and p11, we also recorded neuronal activity during podcast episodes. Patient p6 listened to three stories, each approximately 7 min long, taken from The Moth Radio Hour (<https://themoth.org/podcast>). The stories were *Wild Women and Dancing Queens*, *My Father's Hands* and *Juggling and Jesus*. Each episode consists of a single speaker narrating an autobiographical story. Patient p8 listened to *Why We Should NOT Look for Aliens—The Dark Forest*, an educational video created by the Kurzgesagt group (Kurzgesagt) (<https://www.youtube.com/watch?v=xAUJYP8tnRE>). The selected stories were chosen to be varied, engaging and linguistically rich.

Micro-CT

As the recordings were performed only in tissue planned for resection, we first removed a small cube of tissue around the probe and then proceeded with the remainder of the resection. The cube specimens were processed according to previously described methods⁵⁶. In brief, resected specimens were fixed in 4% PFA for 16 h at 4 °C. They were then stabilized using a modified stability buffer (mStability), containing 4%

Article

acrylamide (Bio-Rad, 1610140), 0.25% (w/v) VA044 (Wako Chemical, 017-19362), 0.05% (w/v) saponin (Millipore-Sigma, 84510) and 0.1% sodium azide (Millipore-Sigma, S2002). The samples were equilibrated in the hydrogel solution for 16 h at 4 °C before undergoing cross-linking at -90 kPa and 37 °C for 3 h. After cross-linking, excess hydrogel solution was removed, and specimens were washed four times with 1× PBS. Next, the samples were immersed in 0.1 N iodine and incubated with gentle agitation for 24 h at room temperature before being embedded in agarose and imaged using a Zeiss Xradia Context micro-CT at 3 μm per voxel resolution. The acquired back-projection images were reconstructed using Scout-and-Scan Reconstructor (Carl Zeiss, v.16.8) and converted to NRRD format using the Harwell Automated Recon Processor (HARP, v.2.4.1)⁵⁷, an open-source, cross-platform application developed in Python. The 3D volumes were analysed, and optical sections were captured using 3D Slicer⁵⁸. All tissue was inspected with a microscope by S.R.H. and her laboratory, and no abnormalities were reported.

Neuronal data processing

Patients did not experience seizures during the surgery (probably due to propofol anaesthesia), so we did not do any seizure-related data-cleaning. The lack of seizures was confirmed by review of the waveform activity by a trained neurologist.

Motion correction. We used previously developed and validated motion estimation and interpolation algorithms to correct for the motion artefacts from brain movement⁵⁹. Motion was estimated via the DREDge software package (Decentralized Registration of Electrophysiology Data software; <https://github.com/evanol/DREDge>) using either a combination of motion traces obtained using raw LFP and/or AP band data, fine-tuned for individual recordings. Motion correction was then implemented using interpolation methods (<https://github.com/williamunoz/InterpolationAfterDREDge>). Both the AP and LFP band data are motion corrected and were used for further preprocessing and analysis steps. If the estimated motion led to no improvement in the spike locations then spike sorting proceeded with the motion correction package built into Kilosort 4 without performing interpolation.

Unit extraction and classification. Automated spike detection and clustering were performed by Kilosort 2.0 if motion correction was already applied using the DREDge algorithm or KiloSort 4.0⁶⁰ if motion correction was not applied separately. Manual curation of spike clusters was performed using the open-source software Phy⁶¹. Unit quality metrics were calculated using SpikeInterface⁶² and were considered single units if they had a d' greater than 1 and fewer than 3% of spikes were violations of a 2 ms interspike interval refractory period.

LFP data. LFP data were bandpass-filtered between 0.1 and 20 Hz and aligned to task events to extract local ERPs. The LFP band amplitude in the specific bands was calculated by first band-pass filtering the raw signal within defined frequency limits (for example, 70–150 for high gamma) and then taking the absolute value of the Hilbert-transformed complex signal. Given the high correlation between adjacent channels, only ten channels equally spanning the length of the probe were used to calculate statistics.

Neuronal data analysis

All analyses were performed using custom MATLAB code unless otherwise noted.

Motion analysis. The motion-corrected location estimates were obtained at a 250 Hz sampling frequency using the DREDge algorithm. This signal was downsampled to 10 Hz. The power spectrum of the calculated motion was then estimated using Welch's overlapped segment averaging estimator for frequencies between 0.1 and 3 Hz.

The amount of motion was defined as the r.m.s.e. of the location trace of the probes centre relative to its average location.

Tone responses. Both single units and multiunits were used for all analyses. A tone responsive neuron was defined as having a statistically significant increase in the average firing rate in the first second after tone onset (shifted by 50 ms to account for neural processing delay to the hippocampus) relative to the preceding 200 ms baseline ($\alpha < 0.05$, Wilcoxon signed-rank test). Visual demonstrations of the peri-stimulus average firing rate were smoothed through a causal Gaussian filter with a s.d. of 150 ms for visualization; however, all statistical analyses were performed on the raw spike count. Response onset latency was computed as the time taken to the peak response. A mixed Gaussian model with two components was then fit to the distribution of latencies. Given the trough between the two peaks at 291 ms and evidence of average oddball response occurring in the first segment, a window of 0–300 ms was used for analysis characterizing tone and oddball selectivity.

Neural tuning. To determine response tuning properties, we modelled trial responses in the peristimulus period using general linear regression models. Neural data in the analysis time window of 0–300 ms was used for tuning analyses. Unit response was defined as the average firing rate, LFP power was defined as the root-mean-squared value of the band-pass-filtered LFP, and gamma power was defined as the average gamma band amplitude. All vectors were z-scored to allow for comparison of the neural response modulation across units/channels. The independent variables were effects-coded for tone type (frequency 1 versus frequency 2), trial type (standard versus oddball) and an interaction term (conjunctive coding). We set the α level at 0.05 to determine whether the β coefficient for the independent variables was statistically significant.

Neuronal population coding. To determine the information content present in the population, a SVM with a linear kernel was trained using tenfold cross validation for 200 iterations. Accuracy for each iteration was defined as the average accuracy across the folds. Significant coding was defined as the distribution of 200 iterations being statistically different from 0.5 (chance). Algorithm validation was performed by shuffling the dataset and demonstrating that it always performed at chance level. Subsampling was performed to avoid performance bias from an unbalanced dataset (that is, more standard trials than oddball trials). To investigate the neuronal population response dynamics for tone and oddball encoding as a function of time, we used sets of sequential trials (50 trials) from each of the two counterbalanced blocks (total of 100 trials). For example, the first set was using trials 1:50 and 181:230, whereas the last set was using trials 101:150 and 281:330. Decoding analyses were also run separately for early versus late trials (first 75 versus last 75 trials within a 150-trial block) for tone and oddball encoding, respectively.

Neuronal response learning dynamics. Next, to determine the neural mechanism underlying statistical learning required for oddball detection, we evaluated single-trial response dynamics across the neuronal population. For each trial, we generated a neuronal response population vector. We then computed the Euclidean distance ($\|\mathbf{u} - \mathbf{v}\|$) and cosine angle ($\text{invcos}(\mathbf{u} \cdot \mathbf{v} / \|\mathbf{u}\| \times \|\mathbf{v}\|)$) between the mean vector across all standard trials and each individual oddball unit vector, evaluating each as a function of the oddball index.

Mixed-effects models. Where applicable, we used mixed effects models to quantify how task conditions affect spike count and other neurophysiological variables while accounting for the hierarchical data structure of multiple subjects, neurons and channels. For analyses of spike count, we summed spikes over equivalent durations across task

conditions and fit a GLME model using a log link function and modeling spike counts as Poisson distributed. For LFP variables, a LME model was used. All analyses used a random effects structure for neurons or channels nested within-participant.

Continuous-rate RNN model

We implemented a continuous-rate RNN and trained it to perform an oddball detection task, closely mirroring the one used for the experimental dataset. The network contains 200 recurrently connected units (80% of which are excitatory and 20% of which are inhibitory units). The network is governed by the following equation:

$$\tau_i(dx_i/dt) = -x_i(t) + W \cdot r(t) + u(t)$$

$$r_i(t) = 1/(1 + e^{-x_i(t)})$$

$$o(t) = W_{\text{out}} \times r(t)$$

where τ_i represents the synaptic decay time constant, $x_i(t)$ indicates the synaptic current variable for neuron i at timepoint t , W is the recurrent connectivity matrix (N by N , that is, 200 by 200), and $u(t)$ is the input data given to the network at timepoint t . u is a 2-by-200 matrix where the first dimension refers to the number of input channels and the second dimension is the total number of timepoints. A firing rate of a unit was estimated by passing the synaptic current variable (x) through a standard logistic sigmoid function. The output (o) of the network was computed as a linear weighted sum of the entire population of units.

In each trial, the network model receives input data mimicking auditory signals. The input consists of two signal streams, each representing a distinct auditory tone (that is, tone A versus tone B; Fig. 3f,g). Only one tone is presented to the network per trial. The model was trained to produce an output signal approaching +1 when tone A was presented and an output signal approaching -1 when tone B was presented. To closely replicate the experimental task design, we used three different sequential contexts during network training. In the first stage, tone A was presented predominantly (80% of trials), followed by an equal distribution of tone A and tone B (50/50) in the second stage. In the third stage, tone B was predominant (80%).

We optimized the network parameters, including recurrent connectivity, readout weights and synaptic decay time constants, using gradient descent via backpropagation through time (BPTT). The network was required to achieve over 95% task accuracy in the current context before a new context was introduced. To evaluate the model's ability to decode both tone identity and oddball context, we performed linear SVM decoding using population activity from the recurrent units. For each decoding analysis, we generated 100 trials for each condition. A linear SVM classifier was trained using 70% of the trials and tested on the remaining 30%, and this procedure was repeated 100 times to estimate decoding accuracy. Separate SVM classifiers were trained for tone identity and oddball context classification.

Neuronal data analysis: natural language stimuli

Natural language stimuli. All of the patients were native English speakers. The podcast played during the task was automatically transcribed using Assembly AI (<https://www.assemblyai.com/>). The transcribed words and corresponding timestamp outputs from Assembly AI were converted to a TextGrid and then loaded into Praat⁶³. The original .wav file was also loaded into Praat and the spectrograms and labels and timestamps were manually checked and corrected to ensure the word onset and offset times were accurate. This process was then repeated by a second reviewer to ensure the validity of the time stamps. The TextGrid output of corrected words and timestamps from Praat was converted to an Excel file and loaded into MATLAB and Python for further analysis.

Natural language statistics. Word frequency was defined based on a corpus of movie subtitles spanning a total of 51 million words²². Words that did not elicit a response during the duration of the word were excluded from this analysis. To compare the relative contributions to firing rate, a linear model was trained to estimate the logarithmic firing rate from the logarithmic duration and corpus frequency of each word. Word surprisal values were calculated using the GPT-2 large model⁶⁴ from the Hugging Face Transformers library⁶⁵, computing the negative log probability of each word conditioned on the preceding context. Specifically, surprisal was defined by the equation

$$\text{surprisal} = -\log P(w_i | w_{i-1}, w_{i-2}, \dots, w_{i-1}, w_i)$$

where $P(w_i)$ refers to the probability of word i given the preceding words.

We used the pre-trained fastText Word2Vec model in MATLAB to extract word embeddings for all words in our dataset^{66,67}. This pre-trained model provides 300-dimensional word embedding vectors, trained on 16 billion tokens from Wikipedia, UMBC webbase corpus and <https://statmt.org>, to capture semantic relationships between words. Notably, Word2Vec is a non-contextual embedder, so all instances of the same word are represented the same. Some surname words, such as Harwood or proper nouns like Applebee's did not have word embeddings and were discarded from the analysis. A simple linear model was trained to predict the firing rate of individual neurons from the semantic matrices using tenfold cross-validation. Accuracy was defined as the correlation between true and predicted firing rates. Words with 0 Hz or above 25 Hz were removed from this analysis. To prevent overfitting, principal component analysis (PCA) was used to reduce the dimensionality of the semantic embedding vectors to account for 30% of the variance before modelling. This threshold was defined as the minimum of the r.m.s.e. of the model that balanced under and overfitting. To predict future or previous words the alignment between words was shifted forwards or backwards, respectively. This relationship was then fit with a piecewise exponential decay

$$r(i) = \beta_0 \times e^{i/\beta_1} \text{ for } i \geq 0$$

$$r(i) = \beta_0 \times e^{-i/\beta_2} \text{ for } i < 0$$

Where β_0 is the amplitude of the correlation at 0 lag, and β_1 and β_2 are equivalent to the time constant of the decay for positive and negative lags, respectively.

Word embedding, semantic clustering and part of speech classification. To identify the natural semantic categories present in our word data, all unique words heard by the participants were clustered into groups using a word-embedding approach^{21,68}. We used the same 300-dimensional embedding from the previous GLM analysis. To compute and visualize semantic clusters, we first used a t -distributed stochastic neighbour embedding algorithm on word embedding values to reduce the dimensionality of each unique word based on their cosine distance to all other words, therefore reflecting their semantic similarity. Words with similar meanings now have similar 2D coordinates. We then applied the k -means clustering algorithm to these 2D word representations and visualized clustered words on a 2D word map (12 clusters). We then manually inspected and assigned a distinct label to each semantic cluster and adjusted clusters for accuracy. For example, words bordering the edges of clusters would sometimes get misgrouped and were manually corrected. The final 12 semantic categories of the words are body parts, places, emotional words, mental words, social words, objects, visual words, numerical words, actions, identity words, function words and proper nouns. Correction for multiple comparisons was performed using the Benjamini-Hochberg approach. A SVM was

Article

trained for each semantic category (versus all other categories) using a radial basis function kernel. Model training and accuracy metrics were weighted to the relative frequency of each group. We used tenfold cross validation and 200 iterations.

To extract part-of-speech (POS) for each word in the dataset, we used an automated pipeline through Stanford CoreNLP, a natural language processing toolkit²⁵. We initialized a CoreNLParser with the 'pos' tag-type, which specializes in POS tagging. The transcript was first segmented into sentences based on punctuation. Each sentence was then tokenized and passed through the CoreNLParser's tagging function. This process leveraged CoreNLP's advanced linguistic models to analyse the context and structure of each sentence, assigning appropriate POS tags to individual words. The 15 POS types were as follows: noun, adjective, numeral, determiner, conjunction, preposition or subordinating conjunction, auxiliary, possessive pronoun, pronoun, adverb, particle, interjection, verb, wh-word and existential. POS types with fewer than 45 words were removed from analysis. A similar SVM was used for POS.

Probe localization

Intraoperative navigation (StealthStation navigation platform, Medtronic) was used to the label probe entry site after it was removed from the brain. RAVE was used to transform patient-specific coordinates into MNI152 average space and plot them on a glass brain with hippocampal segmentation⁵¹.

Ethics statement

Experiments were conducted according to protocol guidelines approved by the Institutional Review Board for Baylor College of Medicine and Affiliated Hospitals (H-50885 for the Neuropixels recordings, and H-18112 for the EMU recordings).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Core data used for analysis have been uploaded to DABI and is available to anyone with a university affiliation (<https://dabi.loni.usc.edu/dsi/U01NS108923>).

Code availability

Code used to implement the computational modelling and analyses in this study is publicly available at GitHub (https://github.com/Nutti-daLab/rnn_oddball).

52. Zhu, H. et al. Semantic axes in the brain support analogical representations. Preprint at *bioRxiv* <https://doi.org/10.64898/2026.01.28.702241> (2026).
53. Chavez, A. G. et al. Mirror manifolds: partially overlapping neural subspaces for speaking and listening. Preprint at *bioRxiv* <https://doi.org/10.1101/2025.09.20.677504> (2025).
54. Yan, X. et al. Shared neural geometries for bilingual semantic representations. Preprint at *bioRxiv* <https://doi.org/10.1101/2025.11.16.688726> (2025).
55. Singh, H. Bispectral index (BIS) monitoring during propofol-induced sedation and anaesthesia. *Eur. J. Anaesthesiol.* **16**, 31–36 (1999).
56. Hsu, C.-W. et al. High resolution imaging of mouse embryos and neonates with X-ray micro-computed tomography. *Curr. Protoc. Mouse Biol.* **9**, e63 (2019).
57. Brown, J. M. et al. A bioimage informatics platform for high-throughput embryo phenotyping. *Brief. Bioinform.* **19**, 41–51 (2018).
58. Fedorov, A. et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magn. Reson. Imaging* **30**, 1323–1341 (2012).
59. Windolf, C. et al. DREDge: robust motion correction for high-density extracellular recordings across species. *Nat. Methods* **22**, 788–800 (2025).
60. Pachitariu, M., Sridhar, S., Pennington, J. & Stringer, C. Spike sorting with Kilosort4. *Nat. Methods* **21**, 914–921 (2024).
61. Rossant, C. et al. Spike sorting for large, dense electrode arrays. *Nat. Neurosci.* **19**, 634–641 (2016).
62. Buccino, A. P. et al. SpikeInterface, a unified framework for spike sorting. *eLife* **9**, e61834 (2020).
63. Boersma, P. Praat: doing phonetics by computer (Praat Org, 2011).
64. Radford, A. et al. *Language models are unsupervised multitask learners* (OpenAI, 2019).
65. Wolf, T. et al. Transformers: state-of-the-art natural language processing. In *Proc. 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations* (eds Liu, Q. & Schlangen, D.) 38–45 (Association for Computational Linguistics, Online, 2020).
66. Mikolov, T., Chen, K., Corrado, G. & Dean, J. Efficient estimation of word representations in vector space. Preprint at <https://arxiv.org/abs/1301.3781> (2013).
67. Joulin, A., Grave, E., Bojanowski, P. & Mikolov, T. Bag of tricks for efficient text classification. In *Proc. 15th Conference of the European Chapter of the Association for Computational Linguistics Vol. 2* (eds Lapata, M. et al.) 427–431 (Association for Computational Linguistics, 2017).
68. Henry, S., Cuffy, C. & McInnes, B. T. Vector representations of multi-word terms for semantic relatedness. *J. Biomed. Inf.* **77**, 111–119 (2018).

Acknowledgements This project was funded in part by U01 NS121472, the McNair Foundation and the Gordon and Mary Cain Pediatric Neurology Research Foundation. This project was supported by the Optical Imaging & Vital Microscopy Core at the Baylor College of Medicine and by the McNair Foundation.

Author contributions K.A.K., S.S., M.F., W.M., S.S.C., A.C.P., N.R.P., A.J.W., Z.W., B.Y.H. and S.A.S. designed the experiment. K.A.K., E.A.M., S.S., M.F., G.P.B., C.-W.H., N.R.P., A.J.W., A.M.G., V.K., A.M., B.Y.H. and S.A.S. collected the data. K.A.K., E.R.C., E.A.M., S.S., M.F., J.A.A., J.L.B., R.K.M., D.M., M.M., W.M., S.S.C., C.-W.H., A.C.P., N.R.P., Z.W., S.R.H., R.K., N.R., B.Y.H. and S.A.S. analysed and interpreted the data. K.A.K. drafted the manuscript and B.Y.H. and S.A.S. revised it.

Competing interests S.A.S. is a consultant for Boston Scientific, Abbott, Koh Young, Neuropace and Zimmer Biomet, and co-founder of Motif Neurotech. The other authors declare no competing interests.

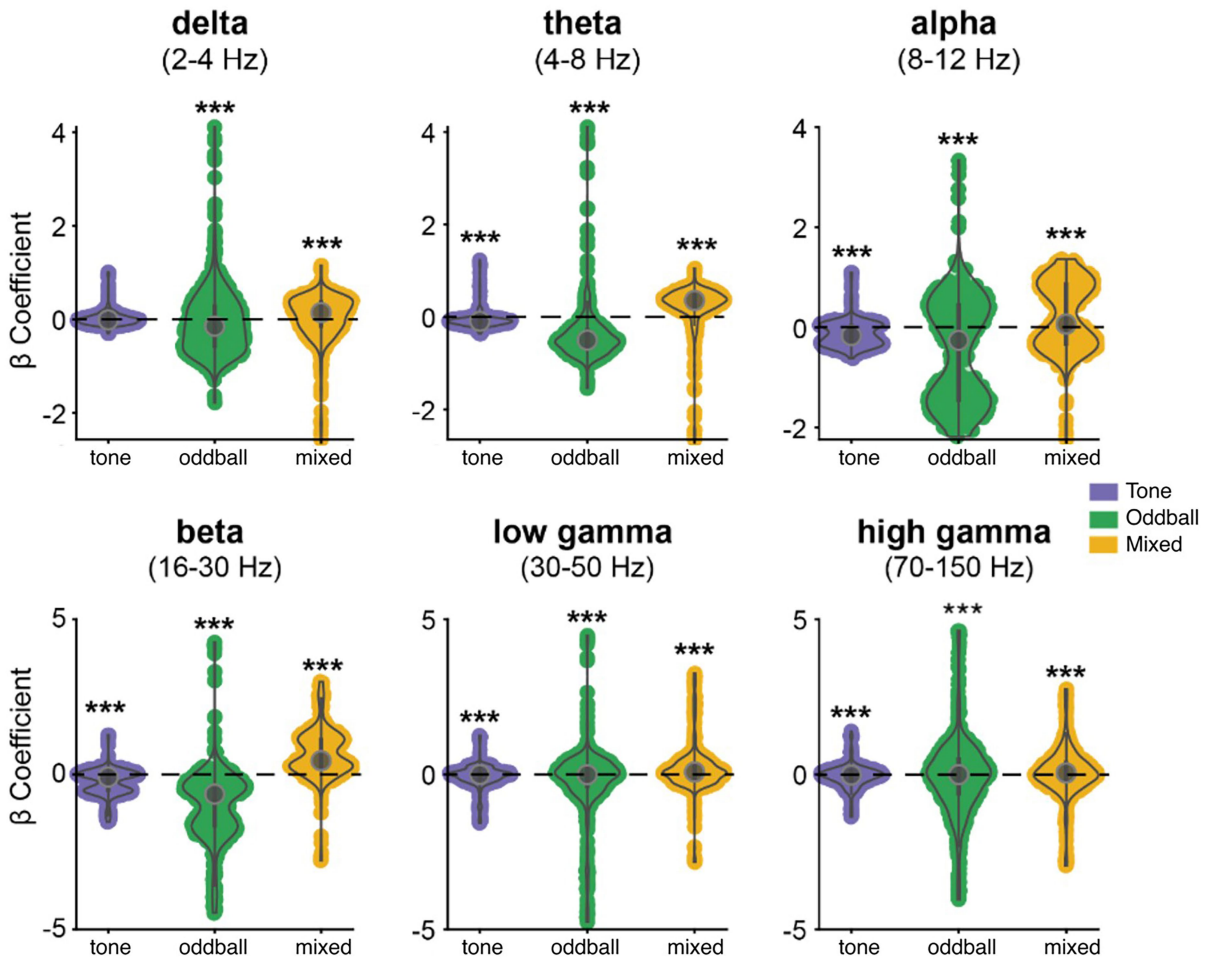
Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41586-026-10448-0>.

Correspondence and requests for materials should be addressed to Sameer A. Sheth.

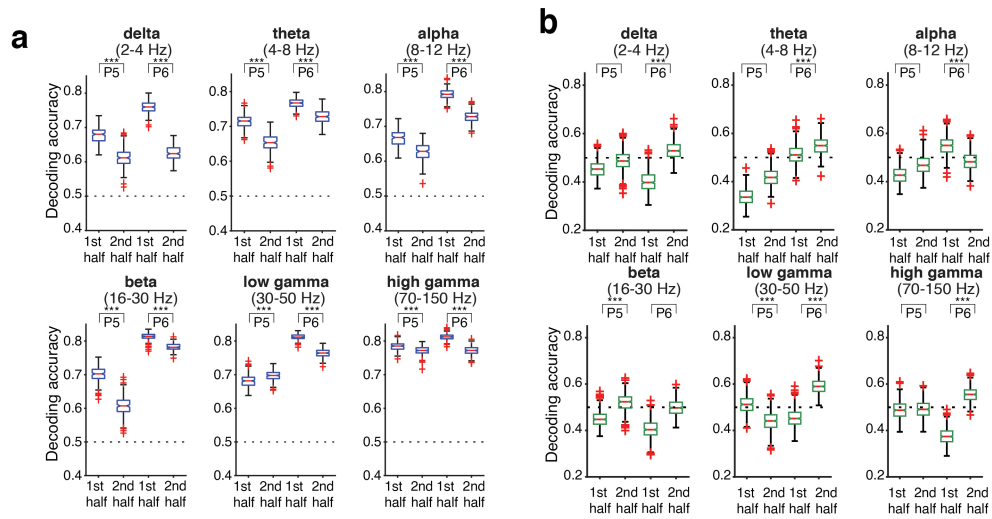
Peer review information *Nature* thanks Robert Knight, Eduardo Sandoval and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at <http://www.nature.com/reprints>.



Extended Data Fig. 1 | Regression models of tone responses as a function of LFP bands. Violin plot showing the distribution of β coefficients obtained from a linear regression model run per channel for each frequency band, to determine response modulation as a function of tone identity (tone β , purple, left) oddball identity (oddball β , green, middle), and an interaction/mixed term

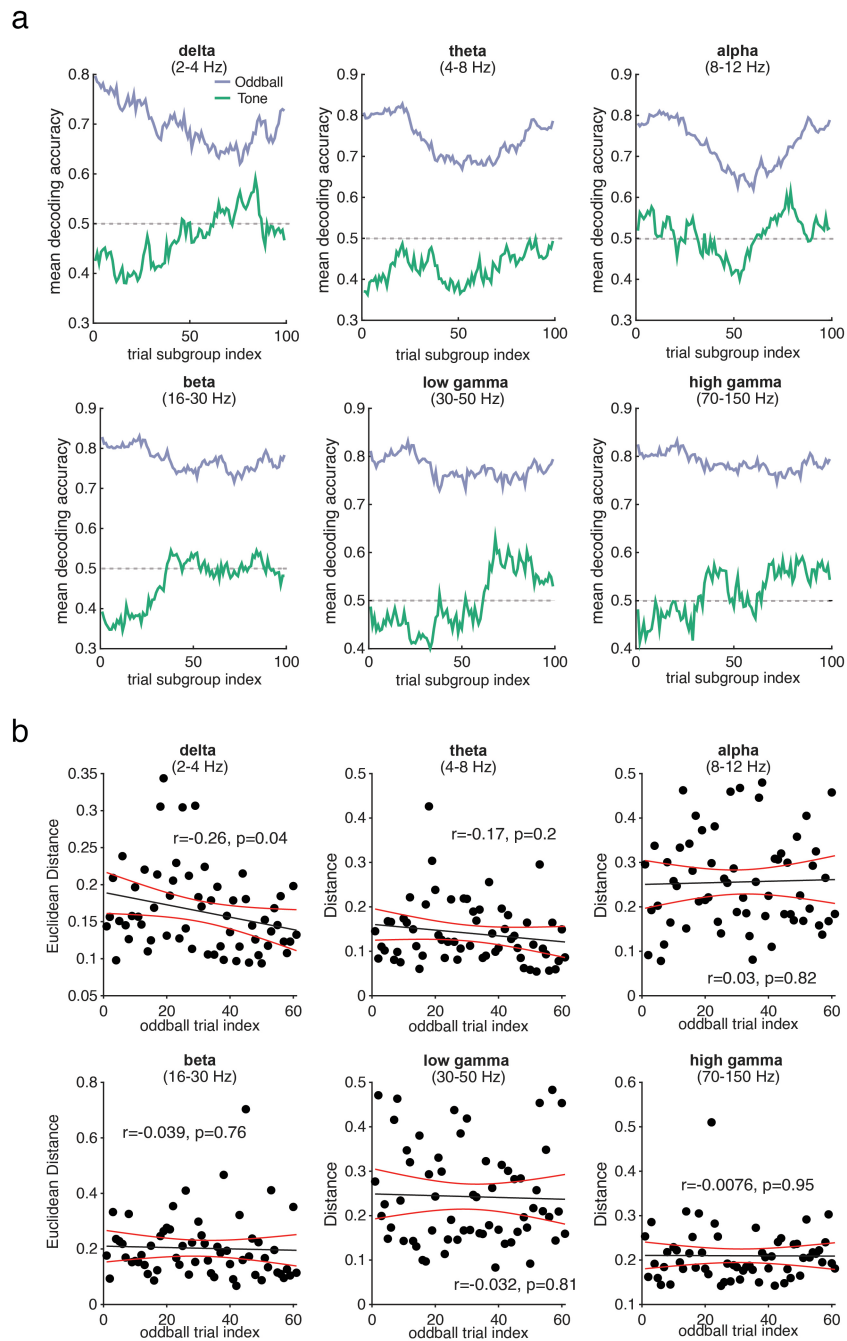
(mixed β , yellow, right). Asterisks reported statistical significance of the difference between the distribution of individual β coefficients and a distribution with a zero median value (nonparametric Wilcoxon's sign rank test). *** denotes p-value < 0.0001.



Extended Data Fig. 2 | Decrease in tone encoding is accompanied by concomitant increase in oddball decoding for multiple LFP bands.

a. Accuracy of tone identity decoding across the population of recorded channels within each pre-defined frequency band for patients P5 and P6, for the first half trials (left) or second half trials (right), combined across both blocks. Statistically significant differences indicated with an asterisk.

*** denotes p-value < 0.0001. **b.** Accuracy of oddball identity decoding across the population of recorded channels within each pre-defined frequency band for patients P5 and P6, for the first half trials (left) or second half trials (right), combined across both blocks. Statistically significant differences indicated with an asterisk. *** denotes p-value < 0.0001.

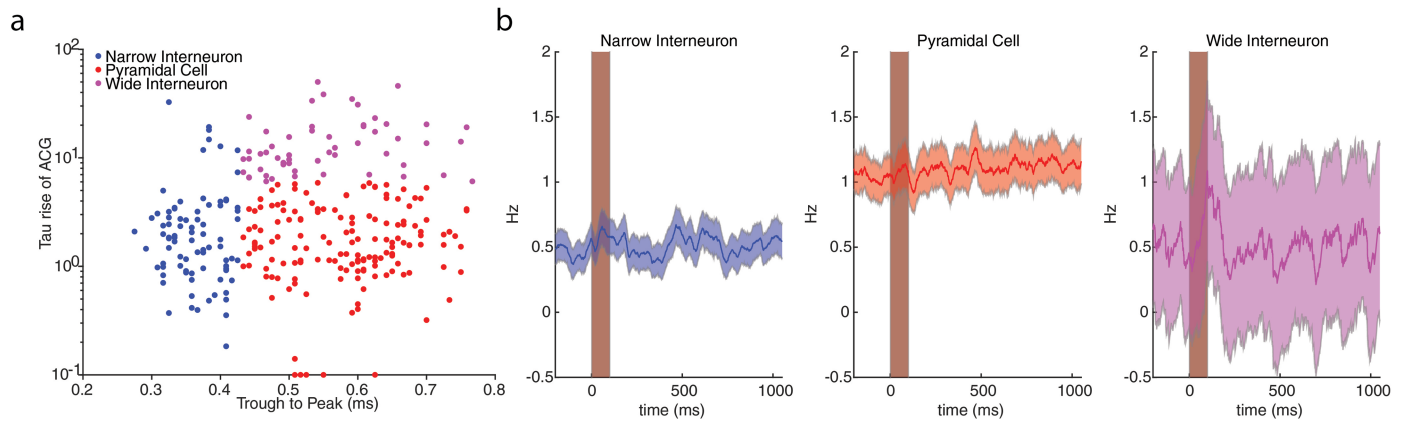


Extended Data Fig. 3 | Evolution of oddball encoding within the LFP.

a. Decoding accuracy as a function of trial position for both patients.

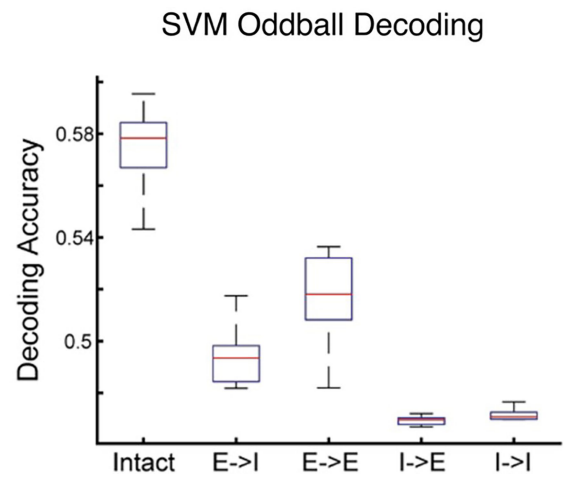
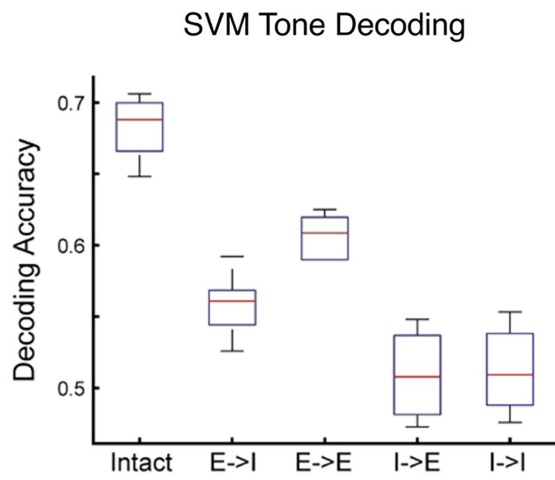
Each point represents SVM accuracy within a set of 50 trials starting at the index location. Decoding accuracy for tone identity is shown in purple, and for oddball identity in green. Dashed line at 0.5 is chance. **b.** Euclidean distance between standard and oddball population response vectors comprising all

channels across both patients ($n = 756$ channels), computed for each oddball trial, separately for responses within distinct frequency bands. Each datapoint (in grey) indicates Euclidean distance per trial, and lines show a linear fit with 95% confidence intervals. Analysis of cosine angle (not shown) demonstrated a highly similar trend.



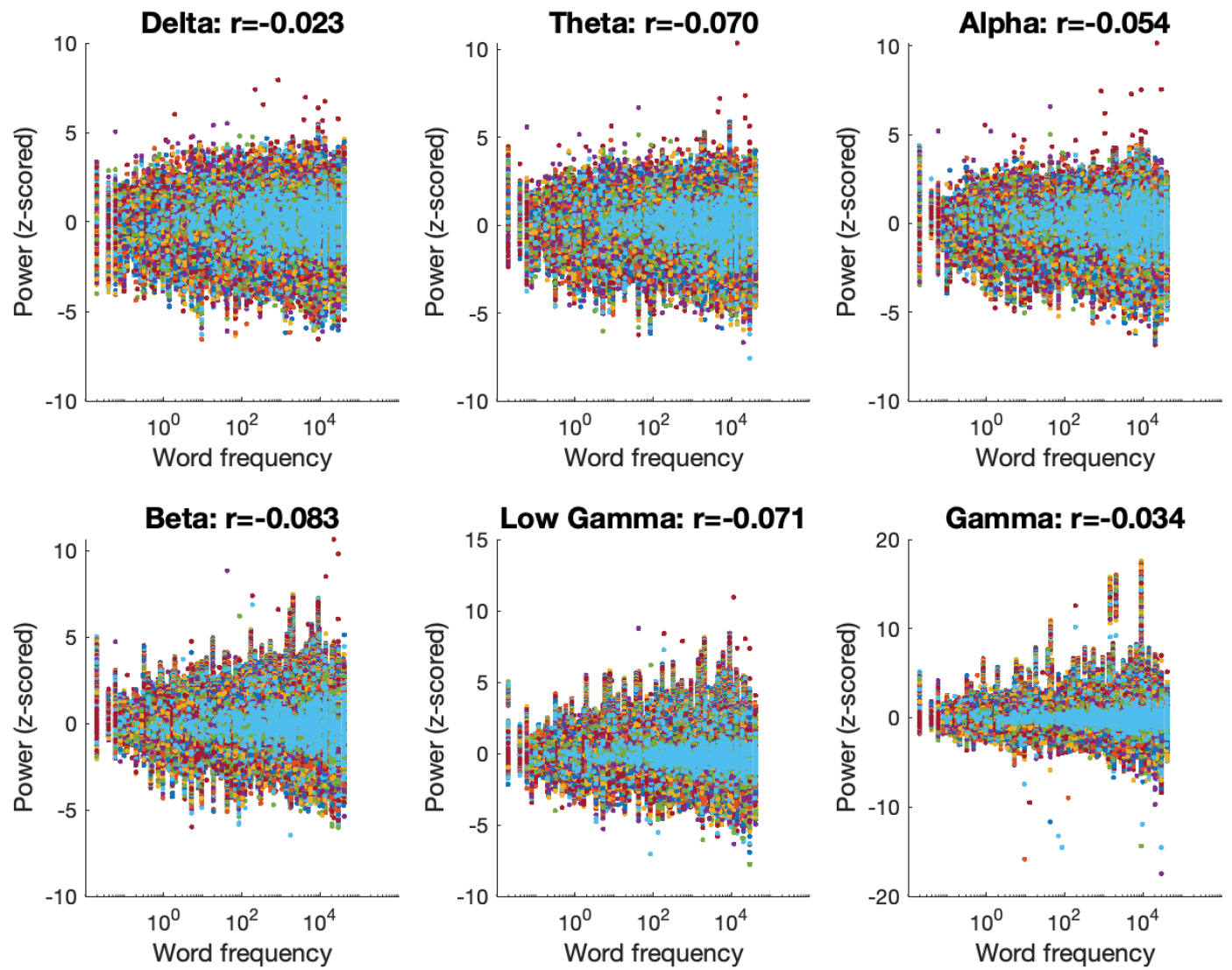
Extended Data Fig. 4 | Categorization of cell types from unit features and tone response by cell type. **a.** The tone response was significant in a larger proportion of interneurons than pyramidal cells (pyramidal cells: 46.7% responders; interneurons: 73.6% responders; $p = 0.0142$, chi-squared test). In the tone/oddball experiment (Figs. 2 and 3), we classified 45.3% of cells as pyramidal cells, 51.2% of cells as narrow interneurons, and 3.5% of cells as wide interneurons using conventional analysis tools (CellExplorer) based on their

spike waveform and autocorrelogram features. In the language analysis (Fig. 4), we classified 65.3% of cells as pyramidal cells, 24.6% of cells as narrow interneurons, and 10% of cells as wide interneurons. **b.** Each subplot shows the firing rate over time (mean \pm SEM across cells) in response to tones. Vertical box: duration of the tone response. Shaded area is the mean \pm SEM across all cells.



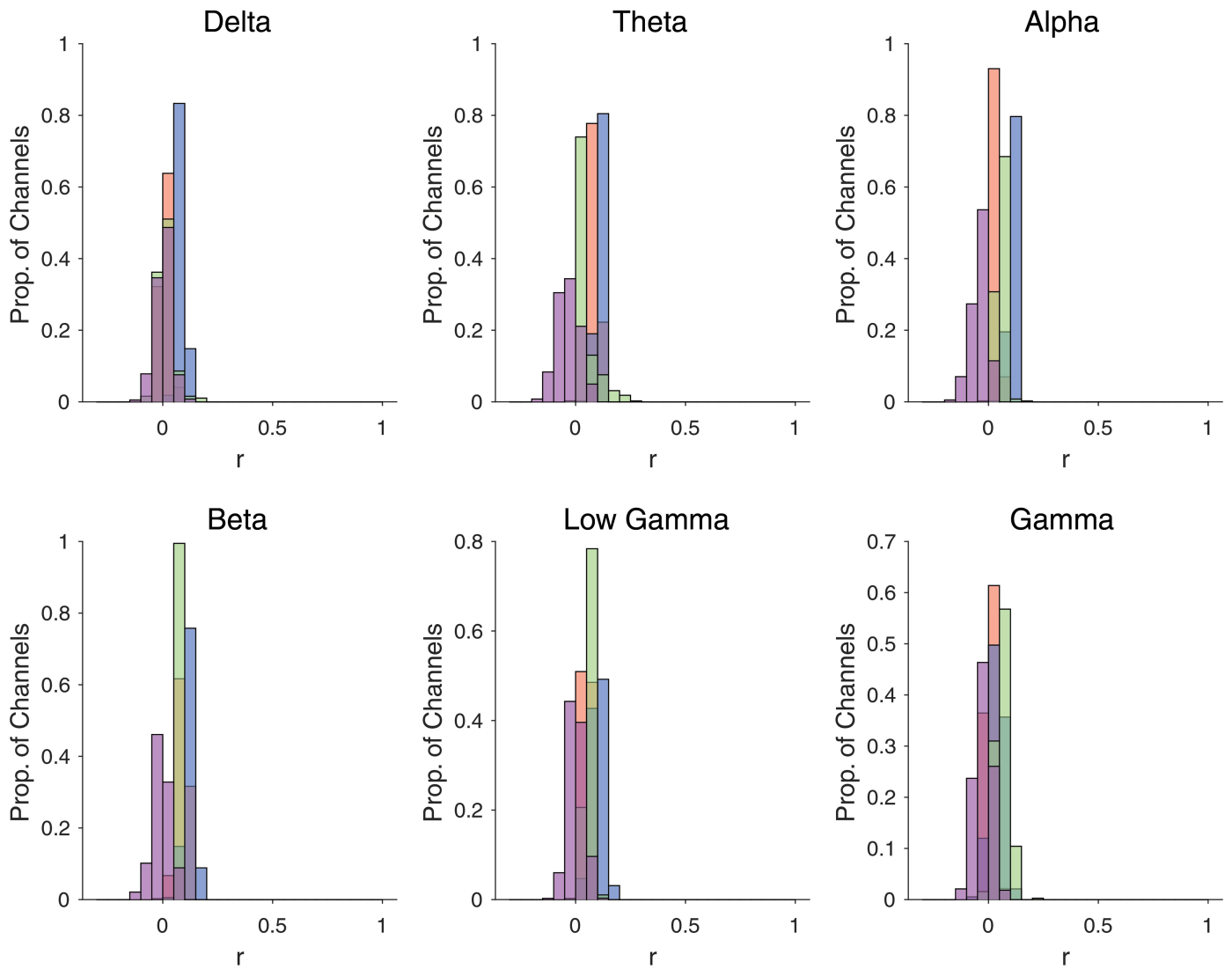
Extended Data Fig. 5 | Inhibitory connections were important for encoding both tone identity and oddball context. Each subtype of recurrent connection in the trained EI-RNN (E->I, E->E, I->E, and I->I) was lesioned by setting the

corresponding weights to zero. We then reran our SVM decoding analysis for Tone (left) and Oddball (right) trials. Box and whiskers are per Fig. 3i.



Extended Data Fig. 6 | Negative correlation between word frequency and evoked power. Z-scored LFP power features were similarly computed for each trial, frequency band, and channel in the 4 patients that participated in the language task. In all 6 bands, power demonstrated a modest but significantly

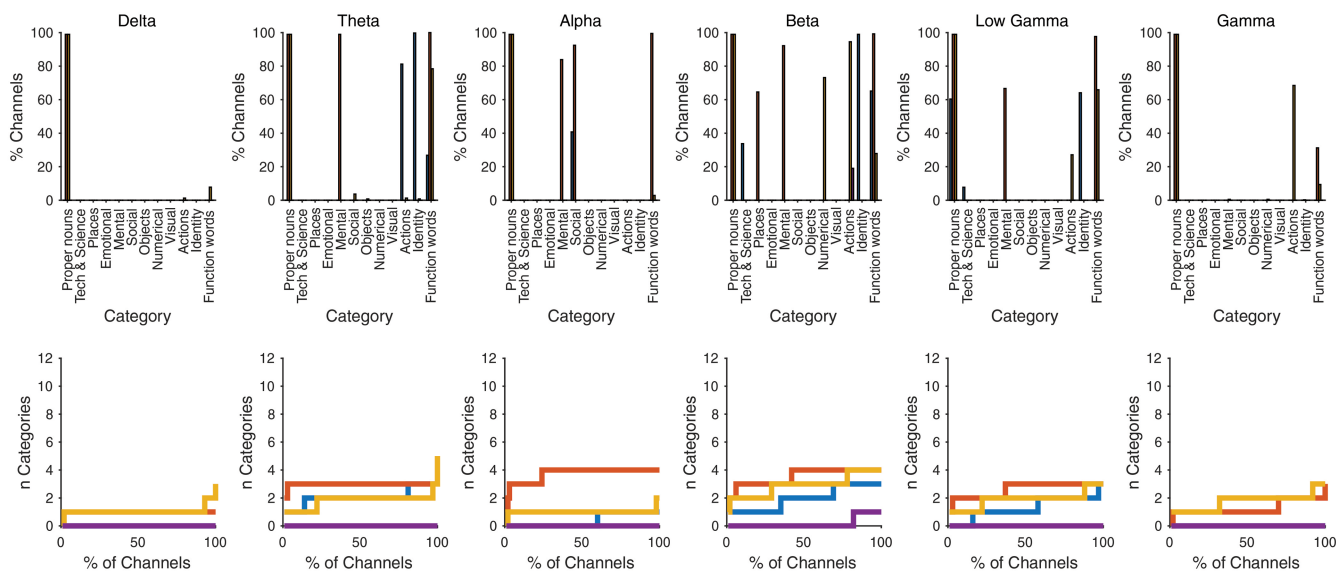
negative correlation with word frequency ($p < 0.0001$; Spearman's correlation, $n = 384$ channels in each of 4 patients), contrasting the results of single unit analysis.



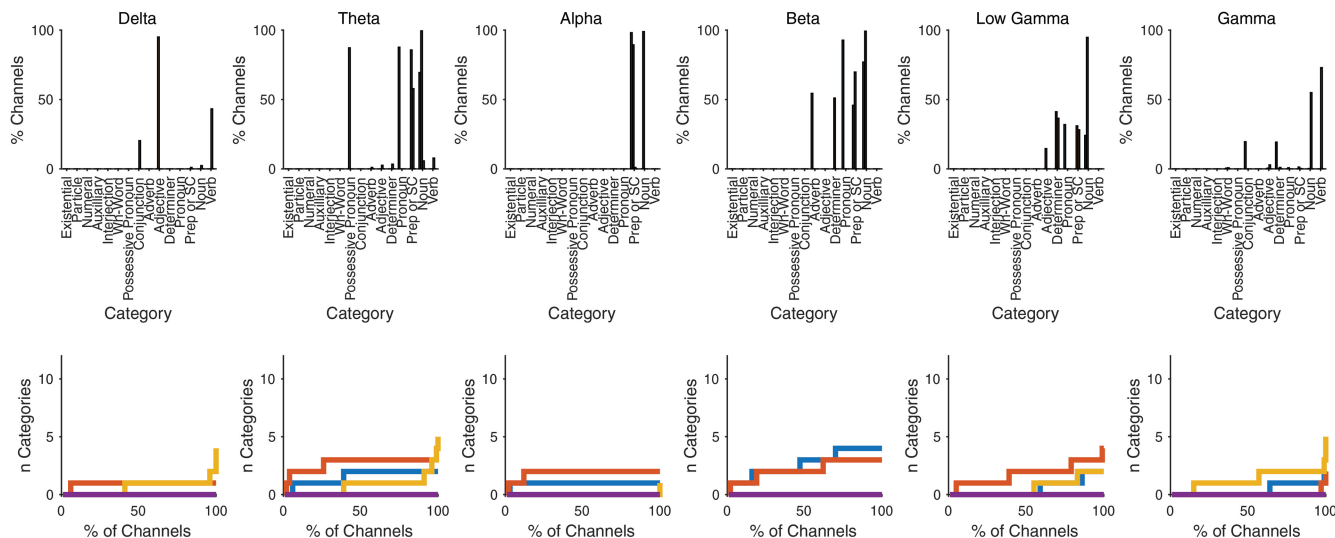
Extended Data Fig. 7 | Semantic embedding prediction performance by LFP power band. Each panel recreates the analysis of Fig. 4c, showing the average correlation between true power and predicted power of a linear model regressing LFP power in the given band vs. the semantic embedding of each word, computed separately for each channel and grouped by patient. Colours indicate distributions for 4 different patients. For all bands, the RMSE of a

linear model significantly outperformed models trained on shuffled data, though with reduced prediction performance relative to single units ($p < 0.05$; delta: mean $R = 0.029$, 46% of channels were significant; theta: mean $R = 0.054$, 75.8% significant channels; alpha: mean $R = 0.039$, 73.7% significant channels; beta: mean $R = 0.067$, 76.4% significant channels; low gamma: mean $R = 0.052$, 76.3% significant channels; gamma: mean $R = 0.021$, 39.3% significant channels).

a

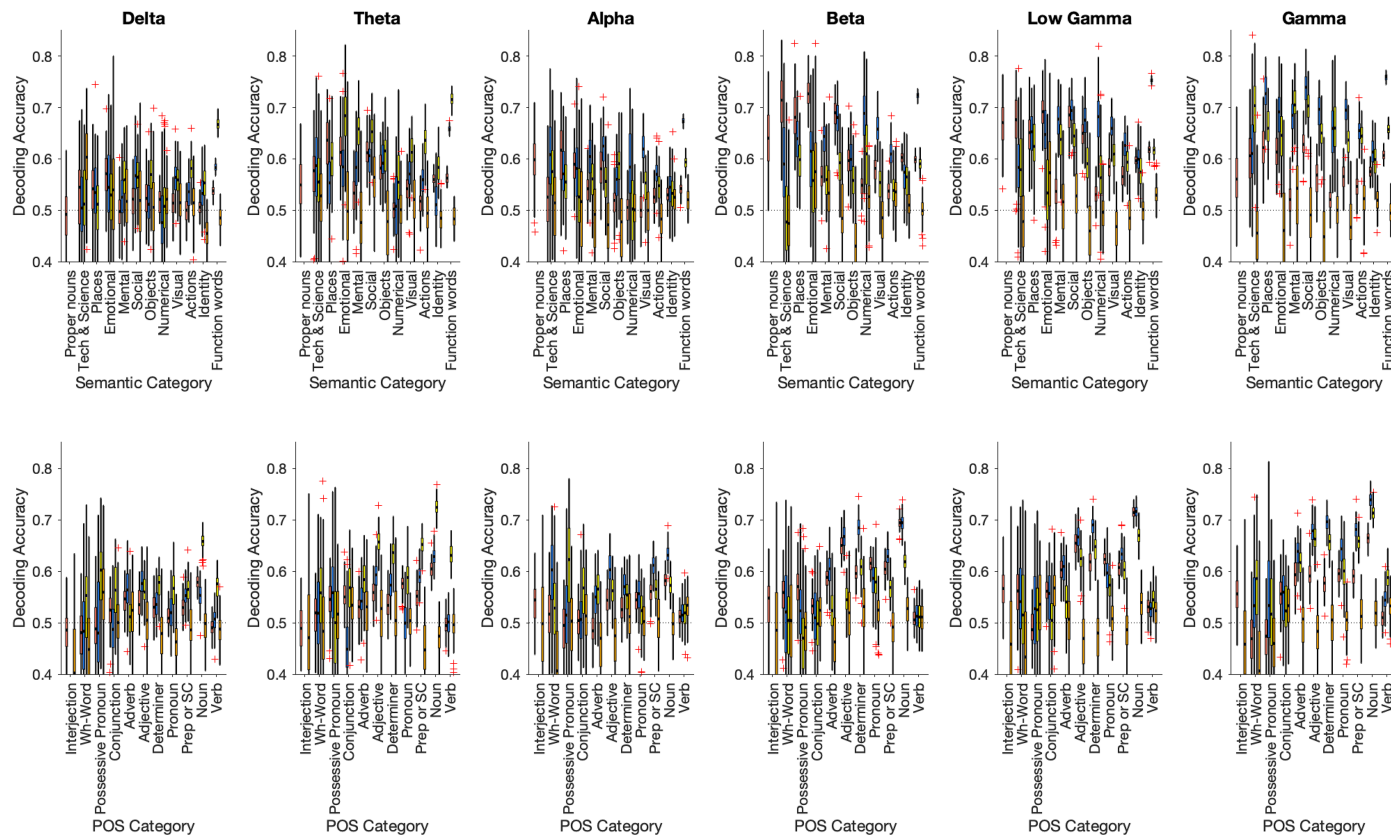


b



Extended Data Fig. 8 | Encoding Analysis of LFP Bands. a. Semantic category prediction performance by LFP power band. Each column recreates the analysis of Fig. 4f and Fig. 4g, quantifying the percentage of channels significantly encoding a specific semantic category vs. other categories. Colours indicate distributions for 4 patients. LFP prediction was overall less significant than single units, with higher variability across patients. Beta and low gamma bands were the strongest predictors of semantic category. Overall, semantic category representation was significant but weaker than that of single units. Category discrimination was greatest in the beta band, followed by alpha, theta, and then low gamma. In the delta band, 50% of channels predicted one category and 2% of channels discriminated between 2 categories. In the theta band, 67% discriminated between 2 categories and 31% discriminated between 3 categories. In the alpha band, 60% of channels predicted 1 category, 26% discriminated between 3, and 19% discriminated between 4 categories. In the beta band, 80% of channels predicted 1 category, 50% discriminated between 3 categories, and 21% discriminated between 4. In the low gamma band, 71% of channels predicted 1 category and 20% of channels discriminated between 3 categories. In the gamma band, 50% of channels predicted 1 category and 25% of channels discriminated between 2. For both types of word feature, category discrimination using LFP was more variable across patients and categories than was observed in the single unit analysis. For example, proper nouns were

discriminated in 100% of channels in the alpha band for one patient but 0% of channels in another. This observation may reflect the lower-dimensional signal content of LFP recordings, as well as their high signal correlation across densely packed channels of the Neuropixel array. **b.** Part of speech prediction performance by LFP power band. Each column recreates the analysis of Fig. 4i and Fig. 4j, quantifying the percentage of channels significantly encoding a specific part of speech vs. other categories. Colours indicate distributions for 4 patients. Part of speech discrimination by LFP power was overall less pronounced than semantic category discrimination but demonstrated comparable trends across bands. Part of speech discrimination was greatest in the beta band, followed by theta, then low gamma, and alpha. In the delta band, 39% of channels predicted one category and 1% of channels discriminated between 2 categories. In the theta band, 42% discriminated between 2 categories and 20% discriminated between 3 categories. In the alpha band, 50% of channels predicted 1 category, and 22% discriminated between 2. In the beta band, 42% of channels discriminated between 2 categories, 23% discriminated between 3 categories, and 8% discriminated between 4. In the low gamma band, 46% of cells predicted 1 category, 24% of cells discriminated between 2 categories, and 6% discriminated between 3. In the gamma band, 32% of cells predicted 1 category and 12% of cells discriminated between 2.



Extended Data Fig. 9 | Decoding Analysis by LFP bands. Each column recreates the analysis of Fig. 4k (top) and Fig. 4l (bottom), quantifying classifier decoding performance for each word category from bandpower values across channels. Colours indicate distributions for 4 patients. LFP prediction was comparable to that of single units. Frequency bands in the alpha range and above were the strongest predictors for semantic category and part of speech. Decoding accuracy was significantly greater than chance (0.5) for all bands, but overall lower than that of single units. Semantic category prediction performance was highest for gamma, followed by low gamma and alpha bands (delta: mean accuracy = 0.520, 72.9% significant categories; theta: mean accuracy = 0.536, 81.3% significant categories; alpha: mean accuracy = 0.555, 77.1% significant categories; beta: mean accuracy = 0.517, 85.4% significant

categories; low gamma: mean accuracy = 0.564, 85.4% significant categories; gamma, mean accuracy = 0.572, 85.4% significant categories). Part of speech prediction performance was also significant, though modestly lower than that of semantic category prediction, mirroring results of the single unit analysis. Decoding accuracy was highest for alpha power, followed by gamma and low gamma bands (delta: mean accuracy = 0.515, 58.3% significant categories; theta: mean accuracy = 0.523, 60.0% significant categories; alpha: mean accuracy = 0.564, 60.0% significant categories; beta: mean accuracy = 0.473, 68.3% significant categories; low gamma: mean accuracy = 0.535, 70.0% significant categories; gamma, mean accuracy = 0.536, 71.7% significant categories).

Article

Extended Data Table 1 | Patient Characteristics

Index	Label	Side	Handed	Sites	Tasks	Anesthetic drips
1	P3	L	L	MTG, HPC	None	Propofol, fentanyl
2	P4	R	R	HPC	Tones	Propofol, sufentanil
3	P5	R	R	HPC	Tones	Propofol, remifentanil
4	P6	L	R	HPC	Tones, Podcasts	Propofol, remifentanil
5	P8	R	R	HPC	Podcasts	Propofol
6	P9	L	R	HPC	Podcasts	Propofol
7	P11	L	L	HPC	Podcasts	Propofol, remifentanil

L: Left, R: Right, MTG: Middle Temporal Gyrus, HPC: Hippocampus.

Reporting Summary

Nature Portfolio 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 Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the 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
- A statement on 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 statistical parameters 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

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Open-source software SpikeGLX 3.0 and OpenEphys version 0.6x were used for data acquisition. The probe was positioned using the ROSA ONE Brain (Zimmer Biomet) robotic arm and lowered using an AlphaOmega (Alpha Omega Engineering) microdrive. MATLAB (R2024b) was used to generate the TTL pulses that synchronized the neural and audio signal capture. Stealth navigation was used to localize the probe entry site (StealthStation, Medtronic)

Data analysis

Custom code written in MATLAB (R2024b) and Python (3.12.9) was used for analysis. Software packages included DREDge (<https://github.com/evanol/DREDge>), Kilosort (<https://github.com/MouseLand/Kilosort>), Phy (<https://github.com/cortex-lab/phy>), and FOOF (<https://foof-tools.github.io/foof/>), and RAVE (<https://github.com/rave-ieeg>)
The computational modeling software has been made publicly available (https://github.com/NuttidaLab/rnn_oddball)

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 and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Core data used for analysis have been uploaded to DABI and is available to anyone with a university affiliation: <https://dabi.loni.usc.edu/dsi/U01NS108923>

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	Summary statistics are reported in the methods. Sex and gender were not considered in the analysis.
Reporting on race, ethnicity, or other socially relevant groupings	Race, ethnicity, and other socially relevant groupings were not considered in the analysis.
Population characteristics	Summary statistics are reported in the methods. This feature was not considered in the analysis.
Recruitment	Anesthetized Neuropixels: All patients diagnosed with drug resistant temporal lobe epilepsy and scheduled to undergo an anteromesial temporal lobectomy for seizure control from 3/1/2023 to 6/1/2025 were offered participation in our study. Awake EMU: All patients diagnosed with drug resistant temporal lobe epilepsy and scheduled to undergo stereotactic encephalography from 3/1/2024 to 6/1/2025 were offered participation in our study. All patients provided written informed consent to participate in the study and were aware that participation was voluntary and would not affect their clinical course. There were no potential biases in the recruitment.
Ethics oversight	Experiments were conducted according to protocol guidelines approved by the Institutional Review Board for Baylor College of Medicine and Affiliated Hospitals, Houston TX (H-50885 for the Neuropixel recordings and H-18112 for the EMU recordings).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

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

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

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

Sample size	No statistical methods were used to predetermine sample size. Patient recruitment was constrained by the availability of eligible neurosurgical cases and limited intraoperative experimental time, and data were collected from all available participants during the study period. Data analysis was performed after completion of data collection. Within patients, units were included based on data quality, and only well-isolated neurons were analyzed. Sample sizes are consistent with, or larger than, most typically reported in human intracranial single-unit electrophysiology studies. Data were analyzed on an individual patient basis with replication across patients.
Data exclusions	One patient was excluded from analysis after analysis post collection demonstrated high amplitude noise was unable to be accounted for and precluded single clustering based on independent analyses from two separate institutions. Within included patients all data were used for analysis.
Replication	10-fold cross validation was used to ensure the reproducibility of all classification and regression analyses.
Randomization	Patients were assigned to maximize the dataset across stimuli. No covariates would be expected between the tasks performed and patient specific features. Additionally, the majority of analyses were performed within patients, with replication across patients demonstrating generalization.
Blinding	Blinding was not possible due to the features of data collection and analysis. However, all methods were equally applied across patients and across neurons, so blinding is not relevant to this study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input checked="" type="checkbox"/>	<input type="checkbox"/> Plants

Methods

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Plants

Seed stocks

Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.

Novel plant genotypes

Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.

Authentication

Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosaicism, off-target gene editing) were examined.