PNAS
www.pnas.org

Supplementary Information for

Hippocampal theta codes for distances in semantic and temporal spaces

Ethan A. Solomon, Bradley C. Lega, Michael R. Sperling, Michael J. Kahana

Michael J. Kahana
Email: kahana@psych.upenn.edu

This PDF file includes:

Supplementary text
Figs. S1 to S6

www.pnas.org/cgi/doi/10.1073/pnas.1906729116

Supporting Information Text

Supplementary Materials and Methods

Human subjects. For behavioral analyses, 189 adult patients with medication-resistant epilepsy underwent a surgical procedure to implant subdural platinum recording contacts on the cortical surface and within the brain parenchyma. Contacts were placed so as to best localize epileptic regions. Data reported were collected at 8 hospitals over 4 years (2015-2018): Thomas Jefferson University Hospital (Philadelphia, PA), University of Texas Southwestern Medical Center (Dallas, TX), Emory University Hospital (Atlanta, GA), Dartmouth-Hitchcock Medical Center (Lebanon, NH), Hospital of the University of Pennsylvania (Philadelphia, PA), Mayo Clinic (Rochester, MN), National Institutes of Health (Bethesda, MD), and Columbia University Hospital (New York, NY). Prior to data collection, our research protocol was approved by the Institutional Review Board at participating hospitals, and informed consent was obtained from each participant. For electrophysiological analyses, a subset of 96 patients with at least one contact placed in the MTL were used.

Free-recall task. Each subject participated in a delayed free-recall task in which they studied a list of words with the intention to commit the items to memory. The task was performed at bedside on a laptop. The recall task consisted of three distinct phases: encoding, delay, and retrieval. During encoding, lists of 12 words were visually presented. Words were selected at random, without replacement, from a pool of high frequency English nouns (http://memory.psych.upenn.edu/WordPools). Word presentation lasted for a duration of 1600 ms, followed by a blank inter-sitmulus interval of 800 to 1200 ms. Before each list, subjects were given a 10-second countdown period during which they passively watch the screen as centrally-placed numbers count down from 10. Presentation of word lists was followed by a 20 second post-encoding delay, during which time subjects performed an arithmetic task during the delay in order to disrupt memory for end-of-list items. Math problems of the form A+B+C=?? were presented to the participant, with values of A, B, and C set to random single digit integers. After the delay, a row of asterisks, accompanied by a 60 Hz auditory tone, was presented for a duration of 300 ms to signal the start of the recall period. Subjects were instructed to recall as many words as possible from the most recent list, in any order, during the 30 second recall period. Vocal responses were digitally recorded and parsed offline using Penn TotalRecall (http://memory.psych.upenn.edu/TotalRecall). Subjects performed up to 25 recall lists in a single session (300 individual words).

Intracranial EEG recordings. iEEG signal was recorded using depth electrodes (contacts spaced 2.2-10 mm apart) using recording systems at each clinical site. iEEG systems included DeltaMed XlTek (Natus), Grass Telefactor, Nihon-Kohden, and custom Medtronic EEG systems. Signals were sampled at 500, 1000, or 1600 Hz, depending on hardware restrictions and considerations of clinical application. Signals recorded at individual electrodes were first referenced to a common contact placed intracranially, on the scalp, or mastoid process. To eliminate potentially confounding large-scale artifacts and noise on the reference channel, we applied a bipolar re-reference for power analyses, and a common average reference of MTL contacts for connectivity analyses. Spectral analyses avoided the 55-65 Hz and 110-130 Hz ranges to mitigate contamination by line noise. As determined by a clinician, any contacts placed in epileptogenic tissue or exhibiting frequent inter-ictal spiking were excluded from all subsequent analyses.

Anatomical localization. To localize contacts to the hippocampus or parahippocampal gyrus, hippocampal subfields and MTL cortices were automatically labeled in a pre-implant, T2-weighted MRI using the automatic segmentation of hippocampal subfields (ASHS) multi-atlas segmentation method (1). Post-implant CT images were coregistered with presurgical T1 and T2 weighted structural scans with Advanced Normalization Tools (2). MTL depth electrodes that were visible on CT scans were then localized within MTL subregions; parahippocampal cortex, perirhinal cortex, and entorhinal cortex were combined to form our PHG label, while CA1, CA3, dentate gyrus, and subiculum were combined to form our hippocampus label. Exposed recording contacts were approximately 1-2mm in diameter and 1-2.5mm in length; the smallest recording contacts used were 0.8mm in diameter and 1.4 mm in length.

Semantic and temporal clustering analyses. In this study, we asked whether theta power during episodic memory retrieval correlated with distances between word items presented in a free-recall task. Distances were measured either as (1) semantic distance, computed from Euclidean distance in word2vec subspaces, or (2) temporal distances, computed from the serial position of the word in a sequentially-presented list.

The procedure to compute semantic distances is outlined in Fig. 1. The word2vec (3) representation of semantic values for each word came from publicly available vectors, trained on a large (approx. 100 billion words) text corpus from Google News (https://code.google.com/archive/p/word2vec/). These vectors represent each word as a point in a 300-dimensional space. To estimate semantic similarity between words, PCA was applied to the 12x300 matrix representing all of the words presented in a given list, to extract between one and 10 principal components. Similarity was taken as the Euclidean distance between points when projected onto the extracted PCA dimensions. (e.g. Fig. 1D-C). Alternatively, PCA was applied to the 300x300 matrix representing all 300 words encountered in a given experimental session ("Session PCA," see Fig. 2B-C). Clustering and neurophysiological analyses were performed using inter-item distances measured in PCA-derived subspaces of varying dimensionality, by extracting different numbers of principal components.

Euclidean distances were computed between each successive pair of words as they were spoken during the free recall period for each list (e.g. Fig. 3A). Distance were converted to a semantic clustering score, reflecting the rank of the distance for reach recall transition, relative to all possible transitions. This was done by finding the percentile rank for the transition between the first pair of recalled words, relative to all possible transitions that could have been made among all words in the list. Next, the just-spoken word is removed as a possibility, and the procedure is repeated through all recall transitions. The result is a score between 0 and 1 for each recall transition, where 1 reflects the closest-possible transition, and 0 reflects the furthest-possible transition. Repeated words were excluded from all behavioral analyses.

In Fig. 2, each subject is assigned a Z-score that captures their overall degree of semantic recall clustering. For each list, a random subset of words is drawn, matched for the actual number recalled on that list, and the clustering scores are recomputed and averaged across lists. This procedure is repeated 250 times, creating a null distribution of clustering scores for each subject. The true average clustering score is compared to the null distribution to generate a Z-score for each subject. Higher Z-scores indicate a greater overall degree of semantic clustering.

As a benchmark, WordNet-derived distances were also used to measure semantic similarity. WordNet is a database enabling the assessment of word similarity via lexical relations (4). Specifically, we used the Wu-Palmer measure of similarity to measure lexical distance between sequentially recalled words. Procedures were otherwise identical to those used for word2vec similarities.

Temporal clustering scores are computed exactly as semantic clustering scores, using the serial position of each word in the 12-item list instead of locations in word2vec-derived spaces. Therefore, a temporal clustering score of 1 indicates a subject recalled two words in sequence that occurred in sequence, while a score of 0 indicates a transition between words that were presented as far apart in time as possible. Z-scores are computed as above, except the null distribution is created by shuffling the true recall order for each list, instead of drawing a random subset of words.

To assess the degree to which clustering scores reflect increased explained variance in our PCA procedure -- as opposed to recall behavior itself -- we used linear regression to remove the effect of marginal explained variance (Fig. 2C). Specifically, for each subject, we took the first-order difference of their clustering versus PCA dimension curve, yielding a new curve that indicates the marginal increase in clustering for each added PCA dimension. We used linear regression to residualize these curves on explained variance for each PCA dimension, with residuals higher than zero indicating behavioral clustering in excess of what could be predicted by explained variance alone.

Correlating spectral power with transition distance. To correlate spectral power with transition distances during episodic memory retrieval, for each subject we extracted 1-second intervals of iEEG from all (bipolar) electrodes placed in the MTL, immediately prior to each recall vocalization onset (Fig. 3B). Repeated words were excluded from all electrophysiological analyses. Spectral power was measured using the multitaper method as implemented in the MNE Python software package (5). For each interval, theta power was taken as the averaged log-transformed power from 4-8 Hz, using a time-bandwith product of 4 and excluding tapers with <90% spectral concentration. Gamma power was taken as the average log power from 30-55 Hz, and high-frequency broadband was taken as the average from 70-150 Hz, with the 110-130 Hz range excluded to mitigate line noise artifact. Power was averaged across all electrodes in each region-of-interest (either hippocampus or PHG), and z-scored across all retrieval events.

As shown in Fig. 3C, z-scored power was Pearson correlated with semantic/temporal transition distances for all recalls in each experimental session, yielding a correlation coefficient for each subject and each session. Only lists with at least three recall transitions (i.e. four recalled words) were included, and any experimental session with fewer than 10 total recall transitions was excluded from all analyses. Correlation coefficients were Fisher z-transformed and averaged across sessions, yielding a final correlation score for each subject, region, and frequency band. Two-tailed 1-sample *t*-tests against 0 were used to determine whether, across subjects, power was significantly correlated with semantic or temporal transition distances (Fig. 4).

Because correlations do not indicate the actual degree of change in power, we also analyzed the association between power and transition distance by binning clustering scores into "short" (above 0.75) and "long" (below 0.25) categories, and taking the average power across all instances in each bin (see Fig. 3D and Fig. 6).

In occasional cases of closely-spaced recalls, 1-second windows of spectral power overlapped with the preceding window (median inter-response time = 1.78 seconds; SI Appendix, Fig. S5-A). Removing recall transitions when overlaps occurred did not appreciably change any main results (SI Appendix, Fig. S5-B).

Connectivity analyses. To analyze the relationship between intra-MTL connectivity and semantic/temporal transition distances during retrieval, we followed as similar approach as with spectral power. Connectivity was measured as the phase locking value (PLV), which reflects the consistency of phase differences between two electrodes either across trials or time (6). In this study, we assessed PLV over time, which yielded a single-trial measure of connectivity between every pair of electrodes in the MTL for each retrieval event. To get a continuous measure of phase over time, we used the Morlet wavelet transform (5 cycles) with 1-second buffers to remove edge effects. PLVs were assessed for the theta and gamma bands separately, and averaged across all electrode pairs that spanned the hippocampus and PHG. We note that, to avoid analyzing connectivity between bipolar pairs which share a common monopolar contact, we used an MTL-average reference for connectivity analyses; i.e. subtracting the average signal across all channels placed anywhere in bilateral MTL. Procedures were otherwise identical to those described in "Correlating spectral power with transition distance" -- hippocampal-PHG PLV was correlated with semantic/temporal transition distance across all retrieval events for each subject, generating a single connectivity-distance correlation measure for each subject. 56 subjects with contacts placed anywhere in the PHG and hippocampus were included in this analysis.

Multiple linear regression analyses. To account for possible confounds between hippocampal theta and semantic or temporal transition scores, we used multiple linear regression to quantify the contribution of hippocampal theta relative to other variables. Specifically, we tested the possible effect of inter-response time (IRT), or the amount of time between successive recalls, and output position (OP), or the order in which recalled words were spoken for each list. To control for the possible relationship between these variables and theta power, we performed multiple linear regression with IRT, OP, and hippocampal theta as predictors of semantic transition distance (SI Appendix, Fig. S3). IRTs were log-transformed and z-scored, and OPs were z-scored, prior to estimating model parameters.

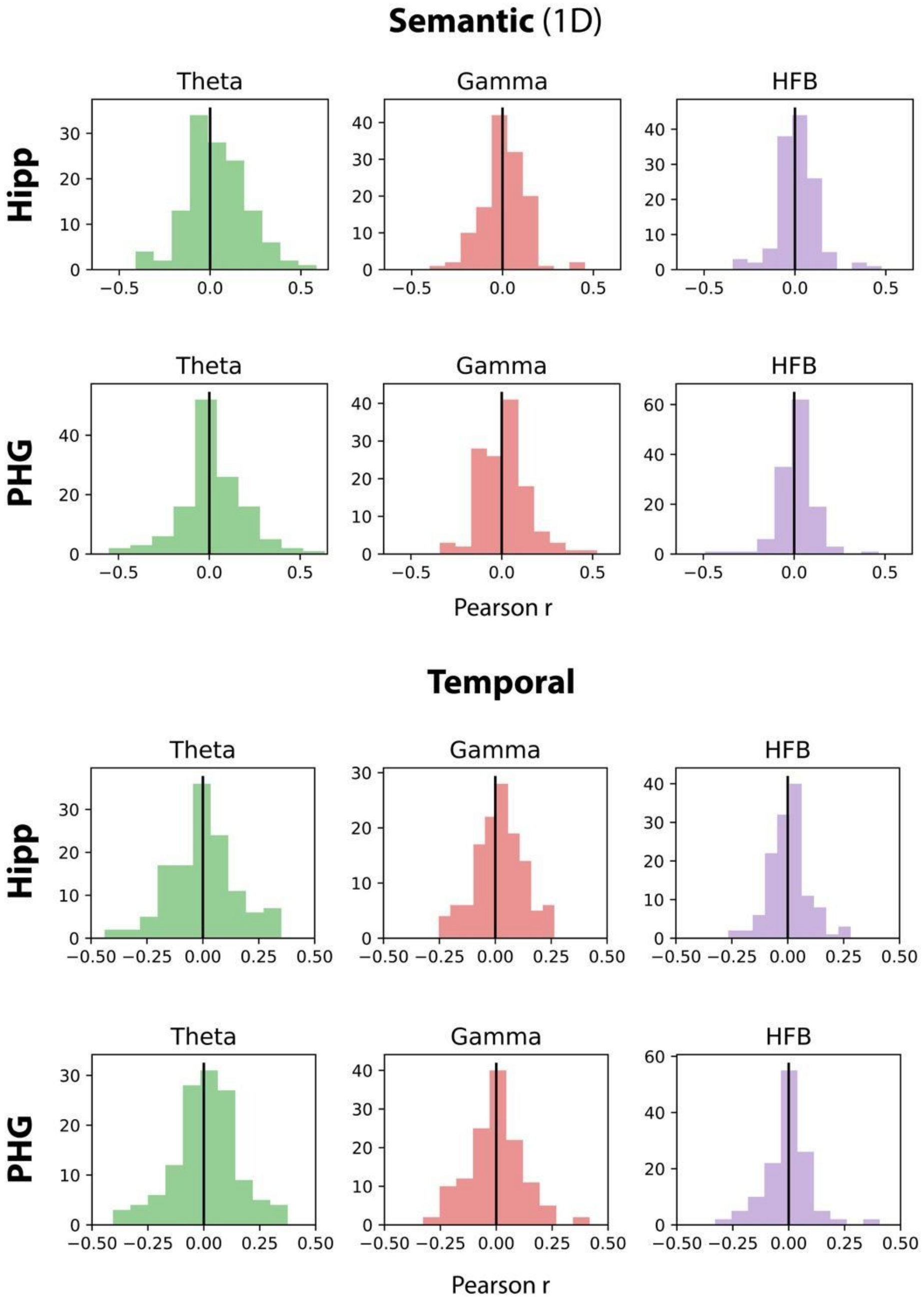

Figure S1. Session-level distribution of subject correlations (Pearson's *r*) for each region-feature-frequency combination (n=128 sessions). Subject-level averages are summarized as bar plots in Figure 4, after applying the Fisher z-transformation. Theta (4-8 Hz), gamma (30-55 Hz), high-frequency broadband (HFB; 70-150 Hz). Only hippocampal theta (top left) was significantly predictive of semantic transition distances in 1-D spaces.

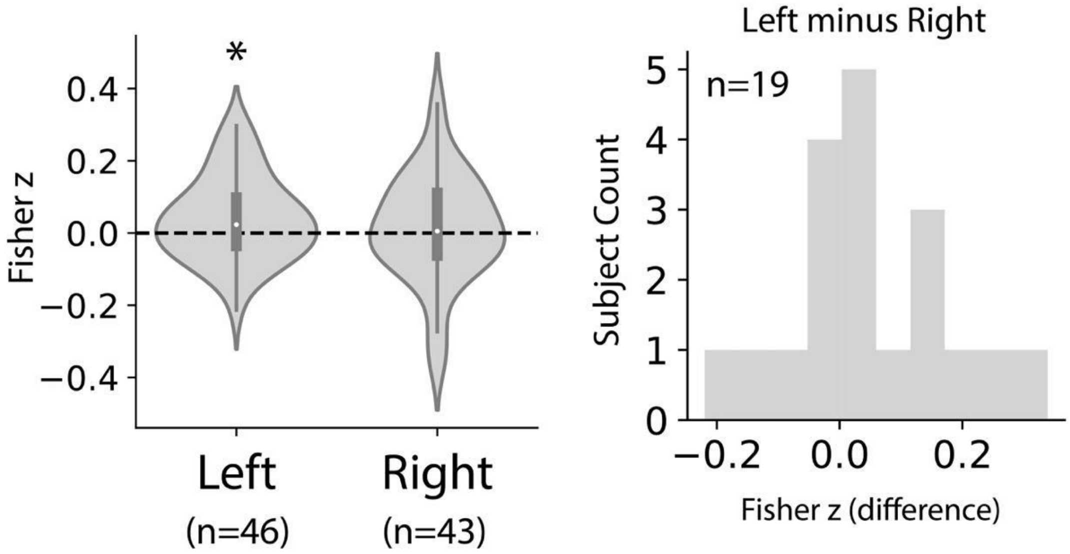

Figure S2. Differences in theta power between left and right hippocampus. *Left:* The correlation between hippocampal theta and semantic transition distance in (A) is driven primarily by activity in the left hippocampus; tested separately, left hippocampus (n=46 subjects) show a significant effect $t(45)=2.33$, $P=0.024$), unlike the right $t(42)=0.60$, $P=0.55$). *Right*: Testing for differences within subjects with bilateral hippocampal coverage, theta-distance correlations are higher in the left than the right, but not significantly so ($t(18)=1.49$, $P=0.15$). $^*P<0.05$.

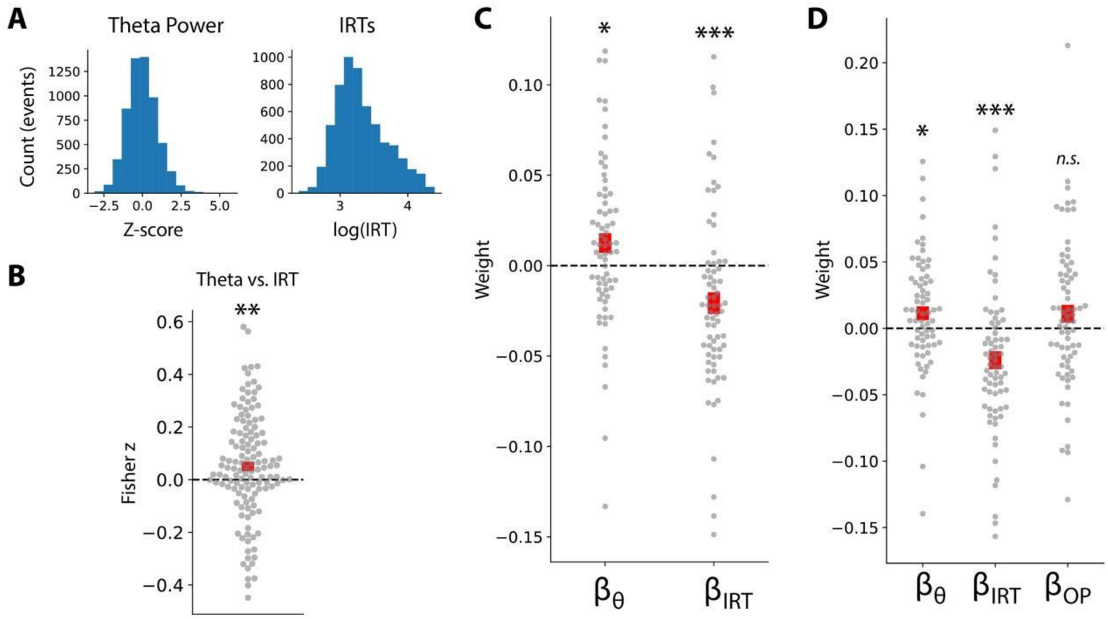

Figure S3. Relationship between hippocampal theta power and inter-response time (IRT). A. Distribution of z-scored hippocampal theta power (*left*) and log-transformed IRTs (*right*) across all events in the dataset (N=5962). **B.** Distribution of (Fisher-transformed) Pearson correlation coefficients between theta power and IRTs for each session (N=128) in the dataset, computed by correlating theta power prior to each recall event with the IRT for the subsequent recall transition, for each session (see Methods). The average Pearson correlation coefficient was 0.05, which was significantly greater than zero (1-sample *t*-test, *P*=0.003). This indicates that theta power is positively predictive of IRT, but the two variables are not colinear and the correlation is weak. **C.** To determine whether theta power is predictive of semantic transition distance independent of IRT, we performed multiple linear regression with IRT and theta power as predictors of semantic transition distance, for each subject and session in the dataset (resulting coefficients were averaged across sessions for each subject; see Methods for details). Across all subjects, the mean beta coefficient for theta power was significantly greater than zero (1-sample *t*-test, *t*(69) = 2.29, *P* = 0.025), recapitulating our earlier result from the main text (main text Figure 4) with the effect of IRT removed. IRTs were negatively correlated with semantic transition distance (*t*(69) = -3.46, *P* = 0.0009), replicating an expected result from prior literature. **D.** Same as (C), with an additional predictor term added to the regression for the output position ("OP") of each recall event. Theta power remained a significant predictor (*t*(69) = 2.18, *P* = 0.032), while output position was nonsignificantly correlated with semantic transition distance across subjects (*t*(69) = 1.66, *P* = 0.10). **P*<0.05, ***P*<0.01, ****P*<0.001, red bars +/- 1 SEM.

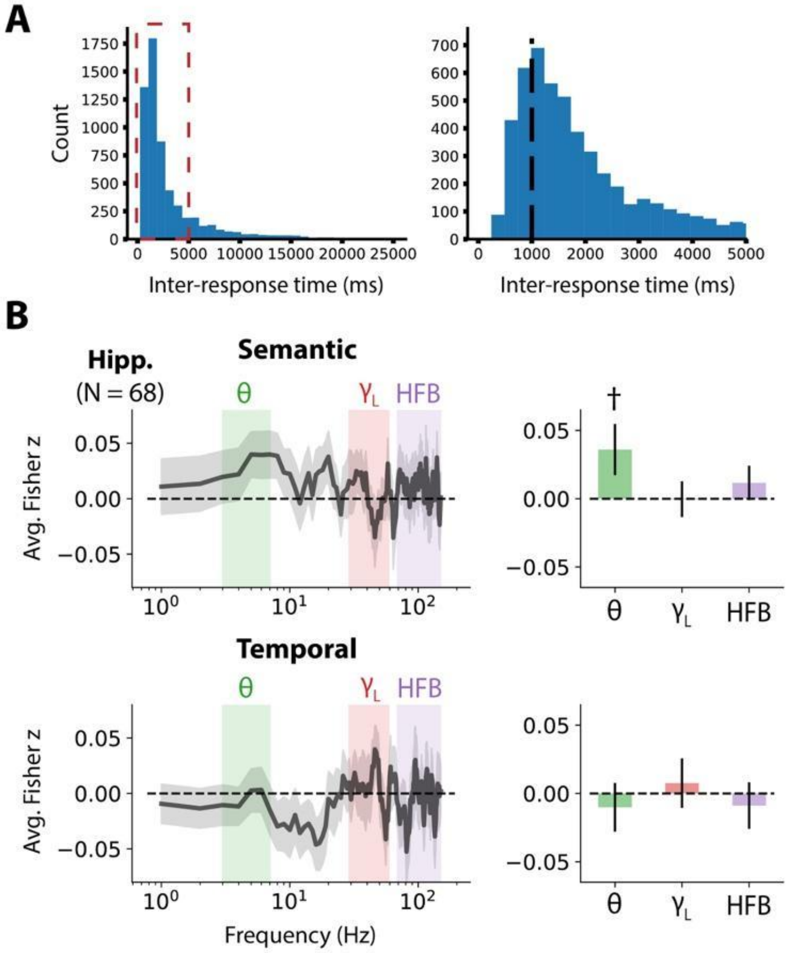

Figure S4. Analysis of spectral power and semantic/temporal transition distance with overlapping 1-second windows removed. A. *Left:* Distributions of inter-response times (IRTs) aggregated across all subjects, indicating the time (in ms) between the onset of successive word vocalizations during free recall. *Right*: Detailed view of highlighted interval on the left. Vertical line drawn at 1 second indicates minimum time to ensure no overlap between windows in which theta power was measured prior to recall. 19.6% of all IRTs were shorter than 1 second; median IRT was 1.78 seconds. **B.** To account for the possibility that 1-second windows of power prior to recall overlap between closely-spaced recalls, all recall transitions were removed from analysis if they occurred within 1 second of the prior recall. In this adjusted dataset, hippocampal theta power remained predictive of semantic transition distance (1-sample *t*-test, *t*(67) = 1.92, *P* = 0.06), while no effect was observed in other frequency bands or the PHG. Note that discarding overlapping events removed 2 subjects from the analysis due to insufficient data. †*P*<0.1, errorbars show +/- 1 SEM.

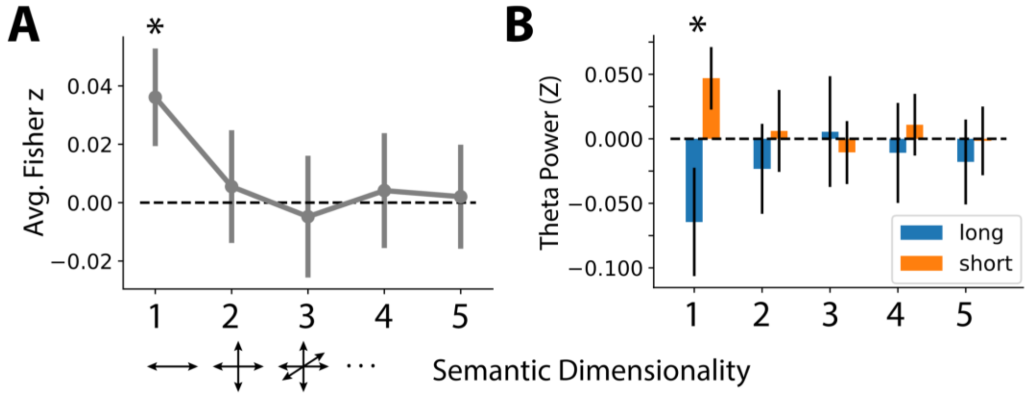

Figure S5. Hippocampal theta predicts semantic transitions only in 1-D semantic space. A. Pre-retrieval hippocampal theta was correlated with subsequent semantic transition distances in PCA-derived subspaces of varying dimensionality. Only distances in 1-D spaces showed a reliable correlation with theta power. **B.** Similar to left, demonstrating the relative change in theta power for long or short transition distances in subspaces of varying dimensionality. **P*<0.05, †*P*<0.1. Error bars show +/- 1 SEM.

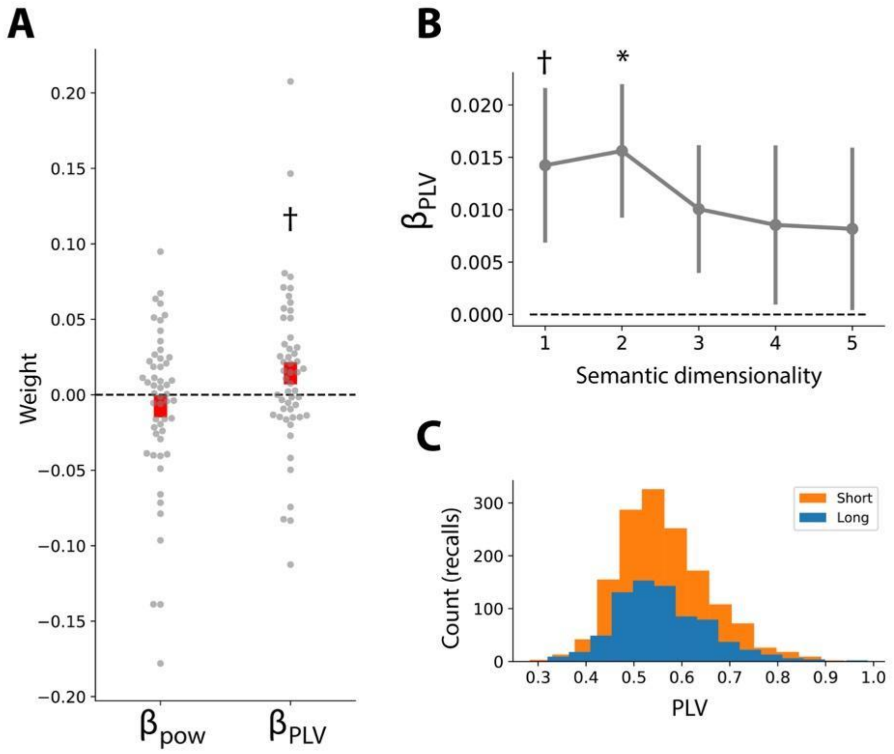

Figure S6. Controlling for spectral power in hippocampal-PHG phase locking. A. Beta coefficients of multiple linear regression with MTL theta power ("pow") and hippocampal-PHG phase-locking ("PLV") as predictors. Controlling for spectral power, population-level phase-locking was marginally significantly correlated with 1-dimensional semantic transition distance (1-sample *t*-test, t=1.93, P=0.06; n=52 subjects). **B.** Mean beta coefficient for hippocampus-PHG theta coupling at varying semantic dimensionalities, controlling for overall MTL theta power; see Methods and (A). Otherwise, same as analysis in Figure 5E. One-dimension, t=1.93, P=0.06; two-dimensions, $t(51)$=2.45, P=0.02. **C.** Distribution of phase-locking values for all events pooled across subjects, binned as "long" (transition distance < 0.25) and "short" (transition distance > 0.75). *P<0.05, †P<0.1, red bars +/- 1 SEM.

References

1. Yushkevich PA, et al. (2015) Automated volumetry and regional thickness analysis of hippocampal subfields and medial temporal cortical structures in mild cognitive impairment. *Hum Brain Mapp* 36(1):258–87.
2. Avants BB, Epstein CL, Grossman M, Gee JC (2008) Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Med Image Anal* 12(1):26–41.
3. Mikolov T, Chen K, Corrado G, Dean J (2013) Efficient Estimation of Word Representations in Vector Space.
4. Miller GA, A. G (1995) WordNet: a lexical database for English. *Commun ACM* 38(11):39–41.
5. Gramfort A, et al. (2014) MNE software for processing MEG and EEG data. *Neuroimage* 86:446–460.
6. Lachaux JP, Rodriguez E, Martinerie J, Varela FJ (1999) Measuring phase synchrony in brain signals. *Hum Brain Mapp* 8(4):194–208.