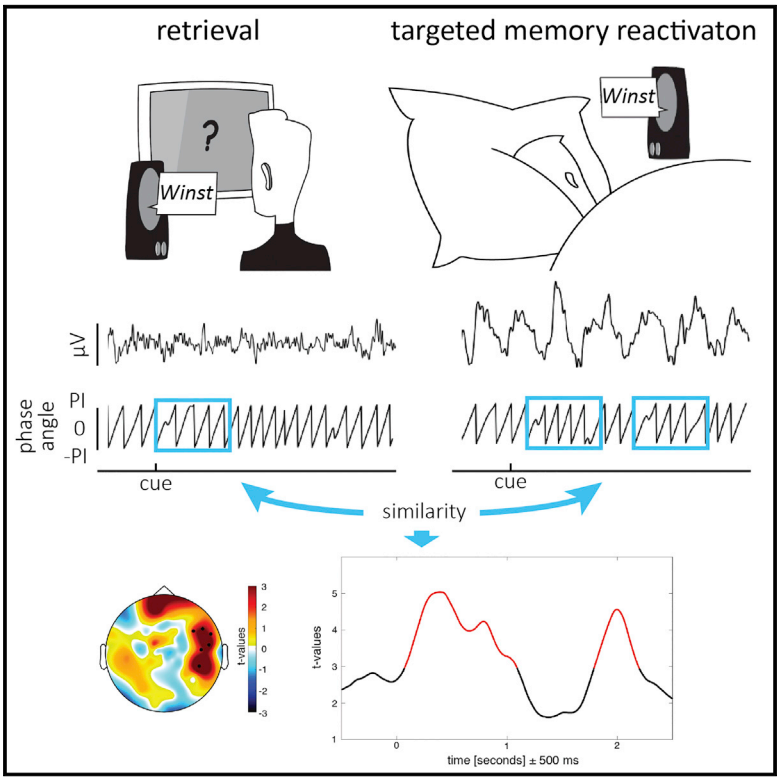
Report

Cell Reports

Theta Phase-Coordinated Memory Reactivation Reoccurs in a Slow-Oscillatory Rhythm during NREM Sleep

Graphical Abstract



Highlights

- Theta orchestrates the reactivation of memories during both wakefulness and sleep
- Reactivation patterns during sleep autonomously re-emerge at a rate of ~1 Hz
- Interrupting the reactivation diminishes the beneficial effects of consolidation

Authors

Thomas Schreiner, Christian F. Doeller, Ole Jensen, Björn Rasch, Tobias Staudigl

Correspondence

t.schreiner@bham.ac.uk (T.S.), tobias.staudigl@cshs.org (T.S.)

In Brief

Schreiner et al. show that cue-triggered memory reactivation shares the same neural signatures during wakefulness and sleep. Theta oscillations orchestrate the reactivation of memories during both physiological states. During sleep, reactivation patterns autonomously re-emerge at a rate of ~1 Hz, indicating a coordination by slow oscillations.

Schreiner et al., 2018, Cell Reports *25*, 296–301
October 9, 2018
https://doi.org/10.1016/j.celrep.2018.09.037

CellPress

OPEN ACCESS
CellPress

Cell Reports
Report

Theta Phase-Coordinated Memory Reactivation Reoccurs in a Slow-Oscillatory Rhythm during NREM Sleep

Thomas Schreiner,[1,2,7,*] Christian F. Doeller,[3,4] Ole Jensen,[1] Björn Rasch,[5] and Tobias Staudigl[2,6,*]
[1]Centre for Human Brain Health, School of Psychology, University of Birmingham, Birmingham, UK
[2]Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, the Netherlands
[3]Kavli Institute for Systems Neuroscience, Centre for Neural Computation, The Egil and Pauline Braathen and Fred Kavli Centre for Cortical Microcircuits, Norwegian University of Science and Technology, NTNU, Trondheim, Norway
[4]Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
[5]Department of Psychology, University of Fribourg, Fribourg, Switzerland
[6]Department of Neurosurgery, Cedars-Sinai Medical Center, Los Angeles, CA, USA
[7]Lead Contact
*Correspondence: t.schreiner@bham.ac.uk (T.S.), tobias.staudigl@cshs.org (T.S.)
https://doi.org/10.1016/j.celrep.2018.09.037

SUMMARY

It has been proposed that sleep's contribution to memory consolidation is to reactivate prior encoded information. To elucidate the neural mechanisms carrying reactivation-related mnemonic information, we investigated whether content-specific memory signatures associated with memory reactivation during wakefulness reoccur during subsequent sleep. We show that theta oscillations orchestrate the reactivation of memories during both wakefulness and sleep. Reactivation patterns during sleep autonomously re-emerged at a rate of ~1 Hz, indicating a coordination by slow oscillatory activity.

INTRODUCTION

The memory function of sleep relies on the reactivation of newly acquired information during non-rapid-eye movement (NREM) sleep (Rasch and Born, 2013). Rodent studies have consistently shown hippocampal reactivation of previous learning experiences during sleep (Chen and Wilson, 2017), and studies in humans have provided first hints indicating similar processes (Peigneux et al., 2004; Schönauer et al., 2017). Furthermore, triggering reactivation processes during sleep by re-exposure to associated memory cues (targeted memory reactivation [TMR]) has been shown to improve memory consolidation (Oudiette and Paller, 2013).

However, the neural mechanisms coordinating reactivation-related mnemonic information in humans remain poorly understood. Furthermore, it is essentially unknown whether memory trace reactivation during wakefulness and sleep is orchestrated by the same neural signatures. Here, we investigated whether memory reactivation during wakefulness and sleep shares oscillatory patterns that carry memory-representation-specific information using electroencephalography (EEG) and multivariate analysis methods. Building on previous findings (Schyns et al., 2011), we hypothesized that low-frequency oscillatory phase conveys a representation (i.e., content)-specific temporal code. We applied a newly developed method (Michelmann et al., 2016) that reveals the phase-related similarity between content-specific memory representations to recently published data (Schreiner et al., 2015; see Figure 1 for experimental design and behavioral results).

We provide evidence for memory-reactivation processes during wakefulness and their reoccurrence during NREM sleep. Theta oscillations orchestrated the reactivation of memories when triggered by memory cues during both physiological states. Reactivation patterns during sleep autonomously re-emerged at a rate of ~1 Hz, suggesting a supra-ordinate coordination by slow oscillatory activity.

RESULTS

Word-Specific Phase Similarity at 5 Hz Indicates Memory Reactivation during Wake Retrieval

First, we aimed at identifying the content specificity of phase and its time course when retrieving the very same memory content during consecutive recall instances, indicating recall-related memory reactivation. The degree of phase similarity for retrieving the same memory content during consecutive recall instances (recall1, recall2) was assessed using the pairwise phase consistency (Vinck et al., 2010) and contrasted between remembered and non-remembered words for frequencies between 3 and 16 Hz (see Figures S1A–S1C for results on the content specificity of our approach). We found significantly higher phase similarity for remembered as compared to non-remembered words in the theta range ($p = 0.006$; corrected for multiple comparisons), peaking at 5 Hz (Figures 2A and 2B). The time course of the phase similarity at 5 Hz displayed an early significant difference between remembered and non-remembered words ($p = 0.008$; corrected for multiple comparisons; Figure 2C; Figures S2A–S2D for unmasked data). Additional analyses indicated that the phase similarity results were not biased by spectral power (see Figures S3L–S3P for details). It seems unlikely that results were driven by similarities in auditory stimulation. Still, we tested this possibility by assessing phase similarity between learning and both recall

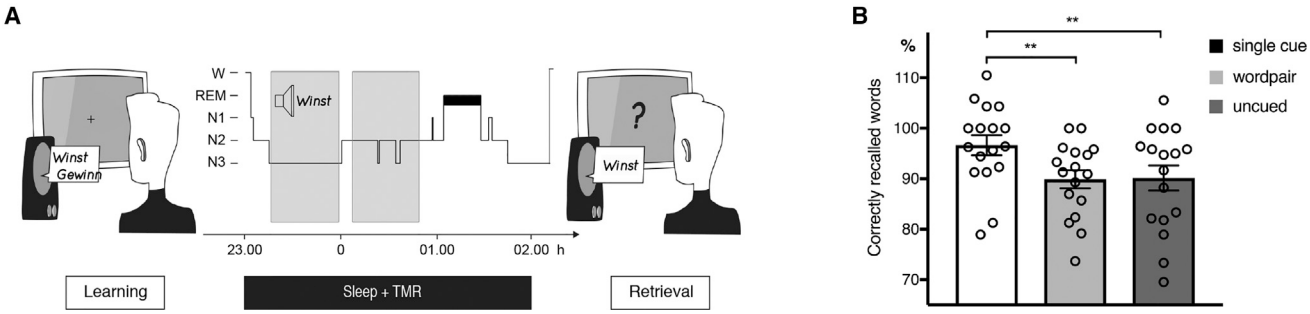


Figure 1. Experimental Design and Behavioral Results

(A) Participants performed a vocabulary-learning task in the evening. They learned to associate Dutch words (cues) with German words (targets). After the initial learning phase, a cued recall, including feedback, was performed (recall1). Afterward, the cued recall was repeated without feedback (recall2). Subsequently, participants slept for 3 hr. During NREM sleep, 80 Dutch words (40 cued and 40 cued + feedback) were repeatedly presented. Memory performance was assessed in the final retrieval phase after sleep

(B) Presenting single Dutch word cues during NREM sleep enhanced memory performance as compared to word-pair TMR and uncued words. Retrieval performance is indicated as percentage of recalled words, with performance before sleep set to 100%.

Values are mean ± SEM. **$p < 0.01$.

instances, with the learning data being segmented around the onset of the Dutch words (thus before any association was learned). Phase similarity was assessed for the very same words and contrasted between remembered and non-remembered words. No significant cluster was observed (both p's > 0.3; see Figures S1F–S1M).

Theta Phase-Coordinated Memory Reactivation Reoccurs during NREM Sleep

The next crucial step was to test whether these content-specific features tracked by phase similarity at 5 Hz would be shared between reactivation processes during wakefulness (recall2) and sleep (TMR). Because memory reactivation during sleep could emerge at any point after TMR cue presentation, phase similarity between recall2 and TMR was examined with a sliding window approach (Michelmann et al., 2016) using the single-trial phase locking value (Lachaux et al., 2000).

Target words remembered after sleep were paired with their equivalent during recall2 and contrasted against non-remembered words. A one-second time window exhibiting the strongest content specificity from the pre-sleep retrieval (center: 0.193 ms; see Figures S3F and S3G for different window lengths) was used as sliding window. Test statistics on the averaged difference between remembered and non-remembered words revealed the reactivation of recall-related phase patterns at 5 Hz during TMR ($p = 0.008$; corrected for multiple comparisons) over right temporal electrodes (Figure 3A; see Figures S1D and S1E for results on the content specificity of our approach and Figures S2E–S2G for unmasked contrasts). No difference in spectral power biased the results (see Figures S3L–S3P for details). To assess the frequency specificity of the obtained results, the same analysis was performed for 3 Hz and 8 Hz (both $p > 0.16$). To test whether our measures were driven by similarity in auditory stimulation, we assessed phase similarity between learning and TMR. No significant cluster was observed ($p > 0.3$; see Figures S1N and S1O).

To examine the time course of the reactivation effect, similarity measures were averaged across significant electrodes and t-statistics were computed for every time point. Two distinct reactivation episodes emerged, peaking at 390 ms ($t_{16} = 4.49$; $p = 0.0003$) and 1,990 ms ($t_{16} = 4.59$; $p = 0.0002$). This pattern of results suggests that presenting a memory cue during sleep triggered re-occurring memory reactivation, fluctuating at a frequency of ~1 Hz (Figure 3D; for analogous analyses using a longer time window, see Figures S3A–S3E).

To test whether slow oscillatory activity might underlie the 1 Hz periodicity found in the TMR similarity measures, we detected slow oscillations (SOs) in all TMR data segments ($n = 581.7 \pm 26.5$). As expected, SOs appeared regularly in the sleep recordings ($n = 416.58 \pm 122.61$), indicating that the fluctuation of phase similarity at a frequency of ~1 Hz might be indeed driven by slow oscillatory activity (for details, see Figures S3H–S3K and Supplemental Experimental Procedures). Testing all combinations of recall and TMR time windows revealed that no additional recall episodes were reactivated during TMR (Figure 3B). To evaluate the sources of the scalp-level effects, phase similarity was assessed on virtual sensors by applying a Dynamic Imaging of Coherent Source (DICS) beamformer. Source-level contrasts exhibited differences in right (para)hippocampal regions as well as more widespread differences in left frontal areas, including the frontal gyrus and insula (Figure 3C).

Our analysis focused on similarity measures between recall and single-cue TMR, because presenting Dutch-German word pairs during sleep abolished the beneficial effects of TMR on later memory performance (Schreiner et al., 2015). Based on this behavioral outcome, we predicted that providing both the cue and target word during TMR should block functionally relevant memory-reactivation processes. There was no significant effect when comparing the averaged difference between subsequently remembered and non-remembered words ($p > 0.17$). As the topographical distribution resembled our main results (Figure 3E), the same electrode cluster was used to characterize the time course. We found an early reactivation episode peaking at 270 ms ($t_{16} = 3.2$; $p = 0.005$) for word pair TMR, thus before the

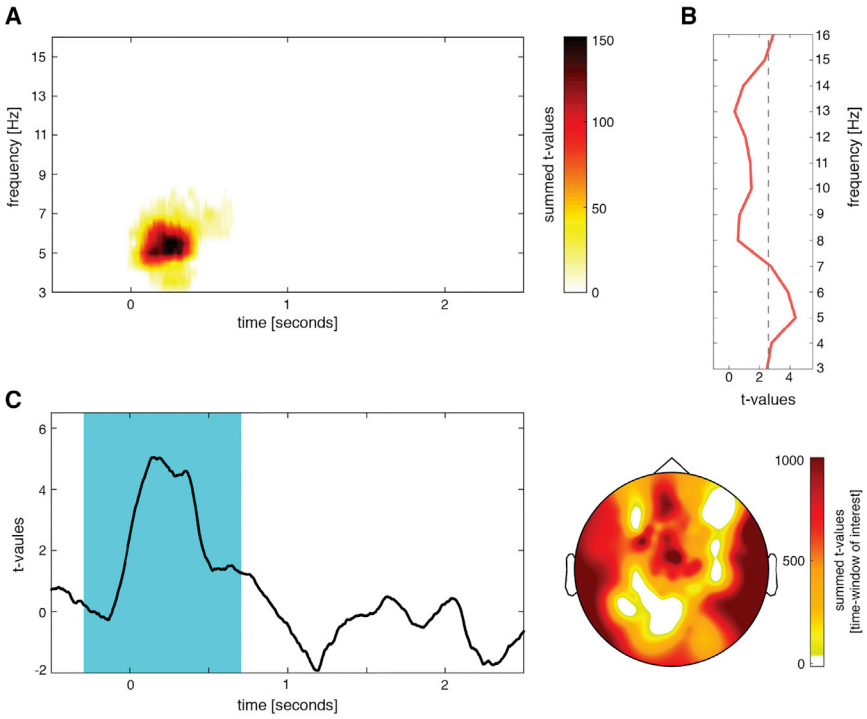


Figure 2. Word-Specific Phase Similarity during Wake Retrieval

(A) Significantly enhanced phase similarity during successful subsequent retrieval was observed early after cue onset (t = 0 s) in the theta range. t-values were summed across electrodes in the significant cluster.

(B) t-statistics of similarity results averaged over time and electrodes indicate a peak at 5 Hz.

(C) Time course and topography of phase similarity at 5 Hz, indicating a rapid reactivation of memory content. The one-second time window around the center of the strongest cluster is highlighted. For the time course, t-values were averaged across all electrodes (n = 83), showing the content-specific phase-similarity effect. The topography displays summed t-values of the averaged difference between 0 and 2.5 s.

See also Figures S1, S2, and S3.

onset of the second word. No later episode was observable, indicating that the presentation of a second stimulus may have blocked further memory reactivation.

DISCUSSION

We show that memory-related reactivation processes during wakefulness and sleep triggered by memory cues share the same neural signature in humans. Theta oscillations at 5 Hz orchestrated the reactivation of memories during both physiological states. A growing number of TMR studies (Laventure et al., 2018; Lehmann et al., 2016; Oyarzún et al., 2017; Schreiner and Rasch, 2015) have already pointed toward a critical role of theta with regards to memory reactivation during sleep, but its exact contribution remained unknown. The current work closes this gap by directly relating memory-associated neural activity during wakefulness and sleep and providing evidence for a common role of theta activity: in both physiological states, the function of theta activity may be to coordinate the reactivation of memories, thus constituting a state-independent feature of memory reactivation.

To investigate the dynamics of reactivation processes in humans, we applied a recently developed method (Michelmann et al., 2016) that detects the phase-related similarity between content-specific memory representations and importantly is robust against variations in the onset of reactivation processes. Here, we show that this procedure constitutes a promising approach, in particular for future research on memory processes acting during offline periods (i.e., rest and sleep).

One of our core findings is that presenting memory cues during sleep triggered re-occurring memory reactivations. After providing a TMR cue, memory-related reactivation patterns re-emerged autonomously at a frequency of ~1 Hz, indicating a coordination by slow oscillatory activity. Models of memory consolidation assume that cortical SOs coordinate reactivation processes as a time-giving pace maker by driving the repeated reactivation of memories in the hippocampus, together with sharp wave ripples and thalamo-cortical sleep spindles (Rasch and Born, 2013). The formation of these spindle-ripple events is thought to be essential for the integration of reactivated hippocampal memory information into neocortical long-term stores (Born and Wilhelm, 2012). Our result that theta-phase coordinated memory reactivation fluctuated at a frequency of ~1 Hz is in line with previous findings, indicating a key role of SOs in guiding reactivation processes (e.g., Johnson et al., 2010). Crucially, our findings expand current models of memory consolidation, as theta activity has not yet been included in theoretical considerations of sleep-dependent memory processing.

Furthermore, our study revealed a re-occurring effect of experimentally induced memory reactivation, as TMR cues triggered a repeated cycling of reactivation patterns during sleep, and presenting a second stimulus abolished this fluctuation and diminished the beneficial effects of TMR. Interestingly, a recent TMR study (Cairney et al., 2018) demonstrated that the decodability of previously learned materials was maximal around 2 s following TMR cues, also hinting toward a perpetuation of the TMR-induced bias. In a similar vein, previous work in rodents has demonstrated that presenting auditory cues during sleep biases the content of associated memory reactivations with maintaining the biasing effect for multiple seconds (Bendor and Wilson, 2012; Rothschild et al., 2017). Although our results indicate that highly comparable processes take place in humans, it is still an open question how the maintenance of TMR induced activity is accomplished.

Our results are also in line with previous findings indicating a role for theta oscillations in mediating communication between

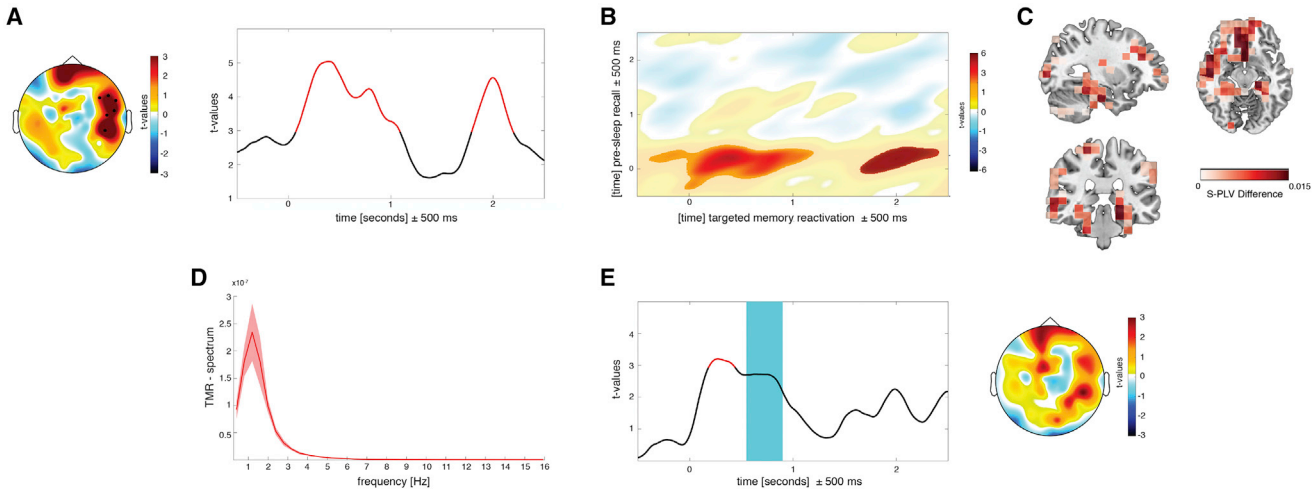


Figure 3. Word-Specific Phase Similarity between recall2 and TMR

(A) Recurrent reactivation of recall-related phase patterns at 5 Hz during TMR emerged over right temporal electrodes. The topography displays the test statistics of the averaged difference in phase similarity between remembered and not-remembered words (0–2.5 s). The time course depicts t-values averaged across highlighted electrodes (n = 6). The phase similarity at a given time point reflects the similarity computed in a window of ±500 ms around this time point.

(B) Assessing phase similarity at 5 Hz between every time point of retrieval and TMR confirmed the re-occurring pattern of similarity.

(C) Source reconstruction. The difference in phase similarity for remembered and not-remembered items indicates effects in right (para)hippocampal regions and left frontal areas.

(D) Frequency spectrum of the TMR similarity measures showed a ~1 Hz periodicity of reactivation processes. Shading denotes SEM.

(E) In line with behavioral predictions, providing a target stimulus after the TMR cue blocked associated reactivation processes. The time course depicts t-values averaged across highlighted electrodes in (A). Presentation of the target word is highlighted in petrol blue. Only a brief reactivation effect at 270 ms (before target word onset) emerged. The topography displays the test statistics of the averaged difference in phase similarity between remembered and not-remembered words (0–2.5 s). No significant cluster was found.

See also Figures S1, S2, and S3.

the medial temporal lobe and neocortical regions (Fuentemilla et al., 2014), possibly conveying hippocampal-driven memory reactivation in the cortex during retrieval (Nyhus and Curran, 2010). Likewise, hippocampal reactivations are thought to drive consolidation processes during sleep, leading to the integration of newly acquired memories into cortical networks (Rasch and Born, 2013). Our source level results corroborate this assumption, as not only right (para)hippocampal areas, previously associated with successful TMR in humans (van Dongen et al., 2012), but also language-related regions in the left frontal cortex (Binder et al., 2009) showed the theta-driven phase similarity effects.

In sum, our results demonstrate that the same neural mechanisms guide memory-trace reactivation during both physiological states of wakefulness and sleep, with a cycling and spontaneous re-processing of memories during sleep when triggered by cueing.

EXPERIMENTAL PROCEDURES

Participants

The data were taken from Schreiner et al. (2015). Thus, detailed information about participants, stimuli, task, data acquisition, and behavioral results can be found in the original article and the Supplemental Experimental Procedures. From the total of 20 participants (13 female; age: 22.45 ± 2.39) who entered the main EEG analyses in the original study (Schreiner et al., 2015), 3 datasets had to be excluded due to extensive artifacts in the pre-sleep EEG data (recall 1/2). Only those 17 participants were included in the illustration of the behavioral results in Figure 1B. The study was approved by the ethics committee of the Department of Psychology, University of Zurich. All subjects gave written informed consent prior to participating.

Word-Specific Phase Similarity during Awake Recall

To detect content (i.e., word) specificity of phase and its time course with regards to successful recall during wakefulness, a modified version of the pairwise phase consistency (PPC) was applied (Michelmann et al., 2016; Vinck et al., 2010). In a first step, oscillatory phase was extracted using complex Morlet wavelets of 6 cycles for all frequencies between 1 and 20 Hz in steps of 0.5 Hz and 1 ms, ranging from 1,000 ms pre-stimulus to 3,000 ms after stimulus onset.

We computed the pairwise phase consistency between the very same words retrieved during consecutive recall instances (i.e., similarity [$word_{a,\ recall1}$, $word_{a,\ recall2}$; $word_{b,\ recall1}$, $word_{b,\ recall2}$; …]). Phase similarity was computed separately per condition ($similarity_{remembered}$ and $similarity_{non\text{-}remembered}$) and contrasted. Retrieval success during recall2 determined the assignment of words to conditions (see Figures S2H–S2J for contrasts determined by memory performance in both recall1 and recall2). Thereby, the degree of phase similarity between identical words and associated memory content was assessed during consecutive recall instances (recall1 and recall2). We assumed that recall processes associated with remembering the very same items should exhibit a higher content-related similarity as compared to non-remembered ones. For each pair of trials, the cosine of the absolute angular distance was then computed and finally averaged across all (remembered and non-remembered, respectively) combinations (Michelmann et al., 2016). A value, representing the average similarity specifically for each set of combinations, was derived for every electrode, frequency, and time bin and subsequently used for statistics. As we assessed phase similarity for the same words presented during recall1 and recall2 and contrasted remembered pairs against an equal number of non-remembered pairs, potential confounding influences of similarity in auditory stimulation should be equal in both conditions and thus controlled for. To further strengthen this point, we assessed the phase similarity between

learning and both recall instances. Importantly, data from learning were also segmented with regards to the onset of the Dutch words. As the German translations were presented with a delay of 3 s, no association-driven reactivation could have happened at this point and the recorded EEG activity primarily mirrors perceptual processing. As power differences can bias phase estimation, we tested whether there was a significant difference in power between conditions.

Word-Specific Phase Similarity between Recall and TMR

We focused our analysis on the single-cue TMR condition. Thus, we tested whether content-specific features of memories would be shared between recall before and single-word TMR during sleep, indicating that TMR during sleep leads to the reactivation of those properties. Phase similarity was assessed for the very same words between recall2 and TMR (i.e., similarity [$word_{a,\ recall2}$, $word_{a,\ TMR}$;…]), separately for each condition (remembered and non-remembered), and contrasted. Retrieval success after TMR determined the assignment of words to condition (see Figure S2K for contrasts being determined by retrieval success in both recall2 and TMR). Phase similarity was contrasted between pairs of successfully acquired memories against an equal number of pairs lacking a stable memory trace. The analysis was restricted to 5 Hz and those electrodes (83 electrodes) that showed the phase similarity effect during wake recall.

We determined phase similarity with the single-trial phase-locking value (S-PLV) (Lachaux et al., 2000; see Supplemental Experimental Procedures). To account for the fact that reactivation processes associated with TMR during sleep might be non-time-locked to the cue, a sliding window approach was utilized. Phase similarity was determined between the 1-s time window from recall2 and each 1-s time window around consecutive time point in the TMR interval (also see Supplemental Experimental Procedures).

The pre-stimulus interval between −500 ms and 0 ms was used as padding to slide the recall window into TMR episodes. This procedure resulted in a single value of phase similarity for every time point and electrode at TMR for any given trial combination. Similarity values thus characterize the similarity of the surrounding 1-s window during TMR to the 1-s time window from recall2. Due to the length of our TMR data segments (ranging to 3 s after stimulus onset) and the 1-s width of the sliding time window (from retrieval), robust similarity values could be obtained up to 2.5 s after cue presentation (for analogous analyses ranging from −500 ms to 3,500 ms, see Figures S3A–S3E).

Next, we explored the time course of reactivation processes during sleep. Electrodes derived from the 5 Hz cluster displaying the significant difference were averaged and subjected to a series of post hoc t tests between remembered and non-remembered combinations for every time point of sleep cueing. We repeated the sliding time window analysis using the same electrodes but varying time windows from the pre-sleep recall. This allowed us to evaluate similarity between every time point of recall and TMR, given an uncertainty of ±500 ms. Afterward, the differences of all combinations were averaged across electrodes. Two control frequencies (3 Hz and 8 Hz) were tested to estimate the frequency specificity.

As TMR during sleep seemed to have triggered reactivation processes in a recurrent fashion, we evaluated whether the similarity measures would fluctuate at a certain frequency ("TMR-spectrum"). We performed a spectral analysis of the time course of the phase similarity differences (for details, see Supplemental Experimental Procedures).

To evaluate whether our similarity measures between recall2 and TMR were influenced by the similarity in perceptual processing, we assessed the phase similarity between learning and TMR. To test whether power differences might have biased the phase estimation, the same control analyses as for the retrieval data were applied. Finally, we tested whether presenting an additional stimulus in the word-pair TMR condition might have interfered with ongoing reactivation processes (Schreiner et al., 2015). Phase similarity was assessed between recall and word-pair TMR in the same way as for the main analysis.

Source Estimation

To estimate the sources of the obtained effects of a virtual electrode approach, we applied the DICS beamforming method (Gross et al., 2001), as implemented in FieldTrip (Oostenveld et al., 2011; see Supplemental Experimental Procedures).

Statistics

Recall-Specific Phase Similarity

Statistical testing of differences in phase similarity between remembered and non-remembered words of recall1 and recall2 was accomplished using a cluster-based nonparametric permutation approach (Maris and Oostenveld, 2007; for details, see Supplemental Experimental Procedures). To estimate the time course and the topographical distribution of the peak frequency, effects at 5 Hz were tested specifically using the same procedure as above against a one-sided distribution (controlling for multiple comparisons in time and space).

Phase Similarity between Recall and TMR

Statistical quantification of the phase similarity between recall2 and TMR contrasting remembered and non-remembered words was accomplished using a cluster-based nonparametric permutation approach. Initially, S-PLV values were averaged over time (0–2.5 s), and differences were tested using paired sampled t tests ($p < 0.05$; two-tailed). To correct for multiple comparisons, 500 permutations were drawn, and the cluster with the largest summed t-value was tested against the permutation distribution. To quantify the temporal characteristics of the obtained effects, phase similarity measures were averaged across the electrodes within a given significant cluster. Paired sampled t tests were computed for every time point ($p < 0.01$; two-tailed).

Cluster-Specific Estimation of Memory Reactivation between Recall and TMR

To statistically test similarity differences between varying time windows from the pre-sleep recall2 and TMR, a series of post hoc t tests was accomplished within the cluster of significant electrodes. t-values, thresholded against a p value of 0.01 (one-sided), were summed up, and 500 permutations were drawn. Following Michelmann et al. (2016), a distribution comprising the strongest cluster, the second strongest cluster, etc. was formed. The obtained clusters were compared against a random cluster distribution. The cluster showing the highest sum of t-values was compared with the distribution of the maximum cluster, and the next cluster was compared to the second strongest cluster, etc. p values were divided by the number of clusters (Bonferroni correction).

DATA AND SOFTWARE AVAILABILITY

All data and analysis codes are available on reasonable request from the corresponding author.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and three figures and can be found with this article online at https://doi.org/10.1016/j.celrep.2018.09.037.

ACKNOWLEDGMENTS

T.S. is supported by a grant of the Swiss National Science Foundation (SNSF) (P300P1_174450). C.F.D.'s research is supported by the Max Planck Society; the Kavli Foundation; the European Research Council (ERC-CoG GEOCOG 724836); the Centre of Excellence scheme of the Research Council of Norway – Centre for Neural Computation (223262/F50); The Egil and Pauline Braathen and Fred Kavli Centre for Cortical Microcircuits; the National Infrastructure scheme of the Research Council of Norway – NORBRAIN (197467/F50); and the Netherlands Organisation for Scientific Research (NWO-Vidi 452-12-009; NWO-Gravitation 024-001-006; NWO-MaGW 406-14-114; NWO-MaGW 406-15-291). B.R. is supported by the SNSF (100014_162388) and the Clinical Research Priority Program (CRPP) "Sleep and Health" from the University of Zurich. O.J.'s research is funded by the Wellcome Trust (207550), the James. S. McDonnell Foundation (220020448), and the Royal Society Wolfson Research Merit Award. We want to thank Maurice Göldi for providing analytical tools used for SO detection.

AUTHOR CONTRIBUTIONS

T. Schreiner and B.R. conceived the design. T. Schreiner collected the data. T. Schreiner and T. Staudigl analyzed the data. T. Schreiner and T. Staudigl wrote

the manuscript. T.S., C.F.D., O.J., B.R., and T. Staudigl discussed the analyses and results and finalized the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: December 12, 2017
Revised: July 17, 2018
Accepted: September 12, 2018
Published: October 9, 2018

REFERENCES

Bendor, D., and Wilson, M.A. (2012). Biasing the content of hippocampal replay during sleep. Nat. Neurosci. *15*, 1439–1444.

Binder, J.R., Desai, R.H., Graves, W.W., and Conant, L.L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. Cereb. Cortex *19*, 2767–2796.

Born, J., and Wilhelm, I. (2012). System consolidation of memory during sleep. Psychol. Res. *76*, 192–203.

Cairney, S.A., Guttesen, A.Á.V., El Marj, N., and Staresina, B.P. (2018). Memory consolidation is linked to spindle-mediated information processing during sleep. Curr. Biol. *28*, 948–954.e4.

Chen, Z., and Wilson, M.A. (2017). Deciphering neural codes of memory during sleep. Trends Neurosci. *40*, 260–275.

van Dongen, E.V., Takashima, A., Barth, M., Zapp, J., Schad, L.R., Paller, K.A., and Fernández, G. (2012). Memory stabilization with targeted reactivation during human slow-wave sleep. Proc. Natl. Acad. Sci. USA *109*, 10575–10580.

Fuentemilla, L., Barnes, G.R., Düzel, E., and Levine, B. (2014). Theta oscillations orchestrate medial temporal lobe and neocortex in remembering autobiographical memories. Neuroimage *85*, 730–737.

Gross, J., Kujala, J., Hamalainen, M., Timmermann, L., Schnitzler, A., and Salmelin, R. (2001). Dynamic imaging of coherent sources: studying neural interactions in the human brain. Proc. Natl. Acad. Sci. USA *98*, 694–699.

Johnson, L.A., Euston, D.R., Tatsuno, M., and McNaughton, B.L. (2010). Stored-trace reactivation in rat prefrontal cortex is correlated with down-to-up state fluctuation density. J. Neurosci. *30*, 2650–2661.

Lachaux, J.-P., Rodriguez, E., Le van Quejen, M., Lutz, A., Martinerie, J., and Varela, F.J. (2000). Studying single-trials of phase synchronous activity in the brain. Int. J. Bifurcat. Chaos *10*, 2429–2439.

Laventure, S., Pinsard, B., Lungu, O., Carrier, J., Fogel, S., Benali, H., Lina, J.-M., Boutin, A., and Doyon, J. (2018). Beyond spindles: interactions between sleep spindles and boundary frequencies during cued reactivation of motor memory representations. Sleep (Basel) *41*, zsy142.

Lehmann, M., Schreiner, T., Seifritz, E., and Rasch, B. (2016). Emotional arousal modulates oscillatory correlates of targeted memory reactivation during NREM, but not REM sleep. Sci. Rep. *6*, 39229.

Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. J. Neurosci. Methods *164*, 177–190.

Michelmann, S., Bowman, H., and Hanslmayr, S. (2016). The temporal signature of memories: identification of a general mechanism for dynamic memory replay in humans. PLoS Biol. *14*, e1002528.

Nyhus, E., and Curran, T. (2010). Functional role of gamma and theta oscillations in episodic memory. Neurosci. Biobehav. Rev. *34*, 1023–1035.

Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J.M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. *2011*, 156869.

Oudiette, D., and Paller, K.A. (2013). Upgrading the sleeping brain with targeted memory reactivation. Trends Cogn. Sci. *17*, 142–149.

Oyarzún, J.P., Morís, J., Luque, D., de Diego-Balaguer, R., and Fuentemilla, L. (2017). Targeted memory reactivation during sleep adaptively promotes the strengthening or weakening of overlapping memories. J. Neurosci *37*, 7748–7758.

Peigneux, P., Laureys, S., Fuchs, S., Collette, F., Perrin, F., Reggers, J., Phillips, C., Degueldre, C., Del Fiore, G., Aerts, J., et al. (2004). Are spatial memories strengthened in the human hippocampus during slow wave sleep? Neuron *44*, 535–545.

Rasch, B., and Born, J. (2013). About sleep's role in memory. Physiol. Rev. *93*, 681–766.

Rothschild, G., Eban, E., and Frank, L.M. (2017). A cortical-hippocampal-cortical loop of information processing during memory consolidation. Nat. Neurosci. *20*, 251–259.

Schönauer, M., Alizadeh, S., Jamalabadi, H., Abraham, A., Pawlizki, A., and Gais, S. (2017). Decoding material-specific memory reprocessing during sleep in humans. Nat. Commun. *8*, 15404.

Schreiner, T., and Rasch, B. (2015). Boosting vocabulary learning by verbal cueing during sleep. Cereb. Cortex *25*, 4169–4179.

Schreiner, T., Lehmann, M., and Rasch, B. (2015). Auditory feedback blocks memory benefits of cueing during sleep. Nat. Commun. *6*, 8729.

Schyns, P.G., Thut, G., and Gross, J. (2011). Cracking the code of oscillatory activity. PLoS Biol. *9*, e1001064.

Vinck, M., van Wingerden, M., Womelsdorf, T., Fries, P., and Pennartz, C.M.A. (2010). The pairwise phase consistency: a bias-free measure of rhythmic neuronal synchronization. Neuroimage *51*, 112–122.

Cell Reports, Volume *25*

Supplemental Information

Theta Phase-Coordinated Memory Reactivation Reoccurs in a Slow-Oscillatory Rhythm during NREM Sleep

Thomas Schreiner, Christian F. Doeller, Ole Jensen, Björn Rasch, and Tobias Staudigl

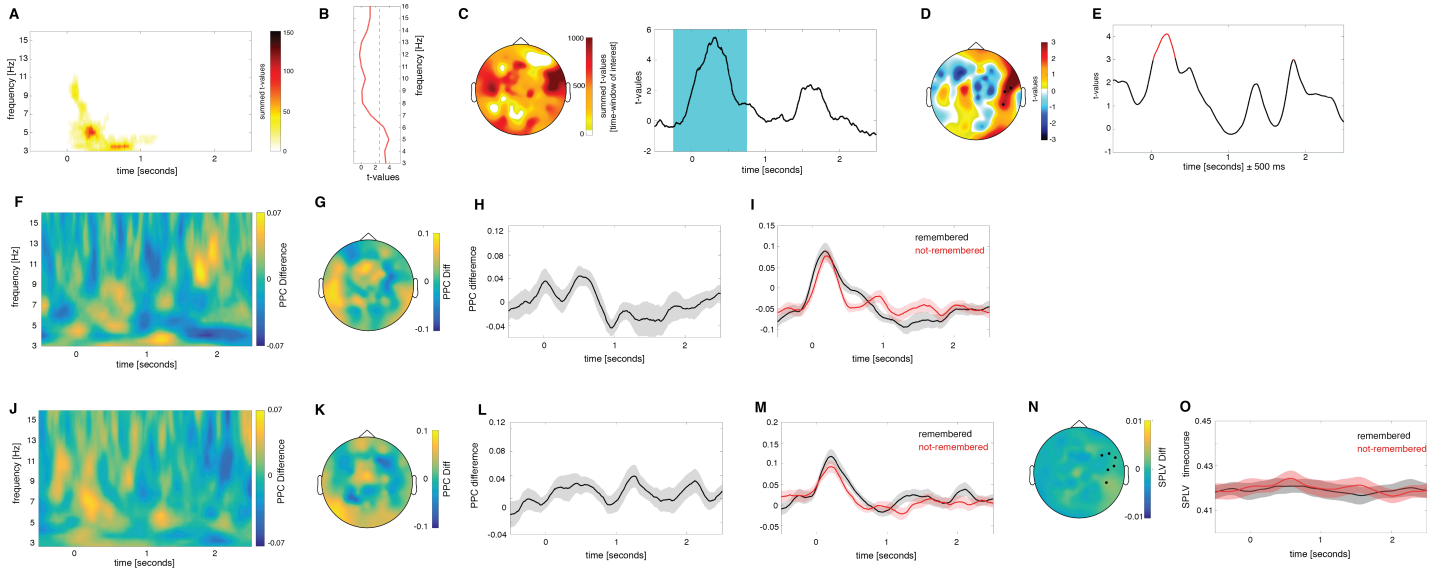


Figure S1 related to Figure 2+3. Assessing content-specificity. (A-E) To evaluate whether phase similarity in the theta range during recall mirrors content specificity, phase similarity was estimated for successfully retrieving the same word in consecutive retrieval instances (i.e. similarity [$word_{a,\ recall1}$, $word_{a,\ recall2}$; $word_{b,\ recall1}$, $word_{b,\ recall2}$; …]) and contrasted against phase similarity for successfully retrieving different words in consecutive retrieval instances (similarity [$word_{a,\ recall1}$, $word_{b,\ recall2}$; $word_{c,\ recall1}$, $word_{d,\ recall2}$; …]). **(A)** We found significantly higher phase similarity when retrieving the same as compared to differing memory content in the theta range (cluster-randomization: P = 0.002, corrected for multiple comparisons), **(B)** peaking at 5 Hz. **(C)** The time-course of the phase similarity at 5 Hz displayed an early significant difference between remembered and non-remembered words (cluster-randomization: P = 0.008, corrected for multiple comparisons). The topography shows summed t-values of the averaged difference at 5 Hz between 0 and 2.5 seconds. These results support the content-specificity of theta phase similarity during recall. **(D+E)** To estimate the content specificity between recall and TMR, phase similarity was estimated for successfully remembered items of the same content (i.e. similarity [$word_{a,\ recall2}$, $word_{a,\ TMR}$; $word_{b,\ recall2}$, $word_{b,\ TMR}$; …]) and contrasted against phase similarity for random pairs of remembered items, thus with differing memory content (i.e. similarity [$word_{a,\ recall2}$, $word_{b,\ TMR}$; $word_{c,\ recall2}$, $word_{d,\ TMR}$; …]) . **(D)** Test statistics on the averaged difference between same and different memory content revealed the reactivation of recall-related phase-patterns at 5 Hz during TMR during sleep (cluster-randomization: P = 0.04, corrected for multiple comparisons) over right temporal electrodes. **(E)** To examine the time-course, t-values were averaged across significant electrodes (n = 6) and t-statistics were computed for every time-point. Two distinct reactivation episodes emerged, peaking at 230ms (t_{16}= 4.06, P = 0.001) and 1800ms (t_{16} = 3.01, P = 0.008). These results further support the content-specificity of theta phase similarity between wakefulness and sleep. **Encoding / Recall phase similarity (unmasked data; (F-M)).** PPC difference between remembered/non-remembered words for the same stimuli during **(F)** encoding and recall1, indicates that similarity measures were not driven by similarities in auditory stimulation **(G)** topography of the phase similarity at 5 Hz **(H)** time-course of phase similarity at 5 Hz (remembered minus not-remembered conditions averaged across the 83 electrodes which showed the content-specific phase similarity effect in the main analysis). **(I)** individual time-courses for phase similarity at 5 Hz for remembered and not-remembered words (again averaged across the 83 electrodes). PPC difference between remembered/non-remembered words for the same stimuli during **(J)** encoding and recall2, indicate that similarity measures were not driven by similarities in auditory stimulation **(K)** topography of the phase similarity at 5 **(L)** time-course of phase similarity at 5 Hz (remembered minus not-remembered conditions, averaged across the 83 electrodes which showed the content-specific phase similarity effect in the main analysis). **(M)** individual time-courses for phase similarity at 5 Hz for remembered and not-remembered words (again averaged across the 83 electrodes). **Encoding / TMR phase similarity (unmasked data; (N+O)).** S-PLV measures contrasted between remembered/non-remembered words for the same stimuli during encoding and TMR at 5 Hz **(N)** topography of phase similarity at 5 Hz indicating no difference between conditions **(O)** time-course of phase similarity at 5 Hz. Shading denotes SEM.

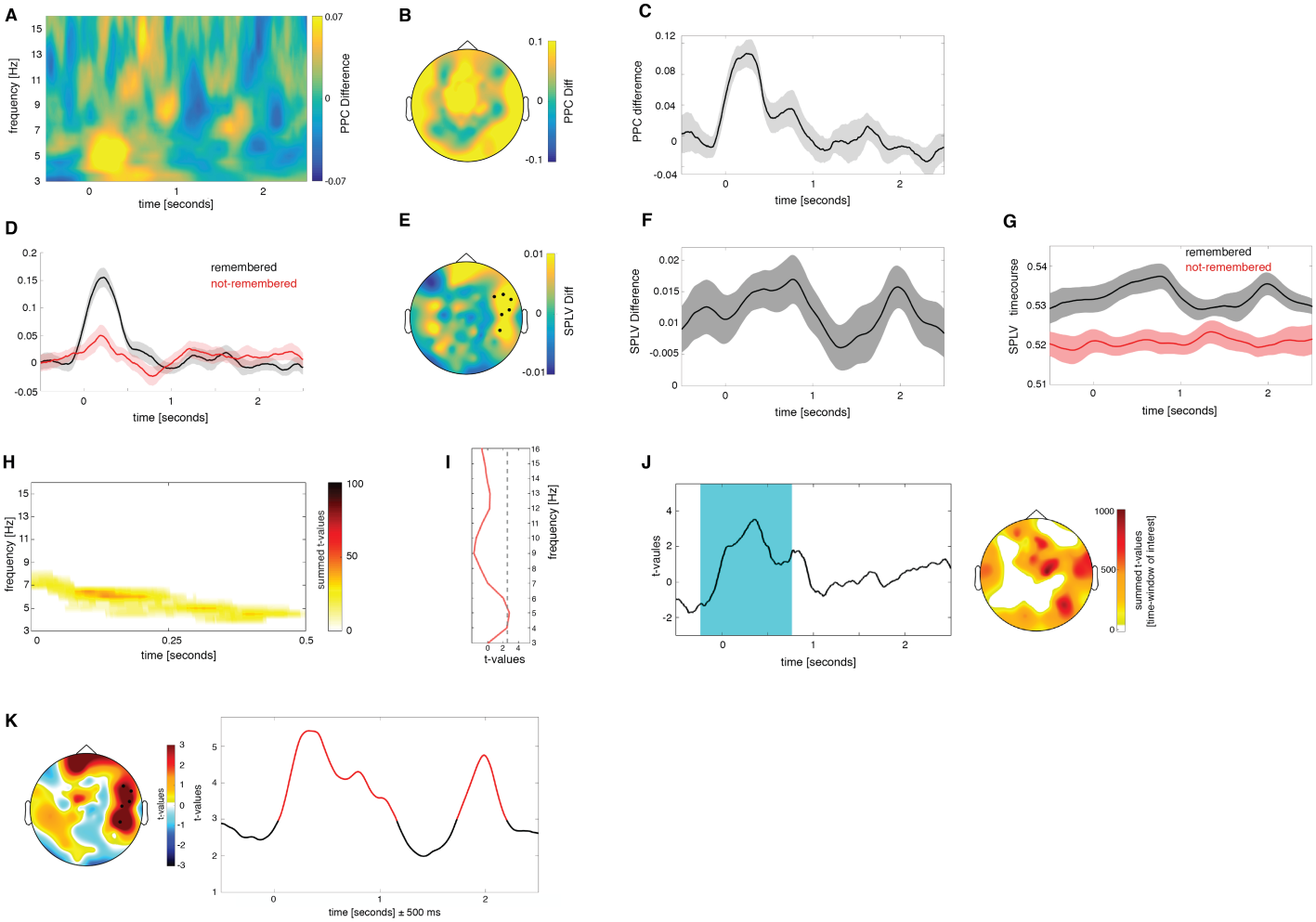


Figure S2 related to Figure 2+3. (A-G) Word specific phase-similarity (unmasked data). (A) PPC difference between remembered/non-remembered words for retrieving the same memory content during consecutive recall instances (recall1 / recall2), shows higher phase similarity in an early time window in the theta range. **(B)** topography of the effect at 5 Hz **(C)** time-course of the difference at 5 Hz (remembered minus not-remembered, averaged across all electrodes (n = 32) showing a significant similarity effect at 5 Hz). **(D)** individual time-courses for phase similarity at 5Hz for remembered and not-remembered words (again averaged across 32 electrodes). **(E)** S-PLV difference between remembered/non-remembered words for the same memory content during recall2 and TMR, indicate a higher difference in phase similarity at 5 Hz over right temporal electrodes. **(F)** time-course of the phase similarity difference at 5 Hz (remembered minus not-remembered conditions, averaged across electrodes highlighted in (E)). **(G)** individual time-courses for phase similarity at 5 Hz for remembered and not-remembered words, averaged across electrodes highlighted in (E). Shading denotes SEM.

(H-J) Phase-similarity during awake retrieval defined by retrieval success in recall1 and recall2 Phase similarity was computed separately per condition ($similarity_{remembered}$, $similarity_{non\text{-}remembered}$) and contrasted. Retrieval success during recall1 and recall2 determined the assignment of words to condition (as compared to the main analysis where only recall2 retrieval success determined assignment of words to condition). Please note that the overall trial number was significantly smaller as compared to our main analysis, due to the fact that in recall1 less words were remembered. Since trial numbers were matched between conditions, also a lower number of trials entered the non-remembered category. Thus, we restricted our analyses to the time-window were we found effects in the main analysis (0-500ms), while including the same frequency range (3-16 Hz). **(H)** We found significantly higher phase similarity for remembered words as compared to non-remembered words in the theta range (cluster-randomization: $P = 0.022$, corrected for multiple comparisons), **(I)** again peaking at 5 Hz (for illustration please see below). **(J)** The time-course of the phase similarity at 5 Hz displayed an early significant difference between remembered and non-remembered words (cluster-randomization: $P = 0.015$, corrected for multiple comparisons). The one second time-window around the center of the strongest cluster is highlighted. For the topography summed t-values of the averaged difference between 0 and 2.5 seconds were plotted. **Phase-similarity between wake/sleep defined by retrieval success in recall2 and TMR (K)** Phase similarity was computed separately per condition ($similarity_{remembered}$, $similarity_{non\text{-}remembered}$) and contrasted. Retrieval success during recall2 and TMR determined the assignment of words to condition (as compared to the main analysis where only retrieval success after sleep determined assignment of words to condition). Test statistics on the averaged difference between remembered / non-remembered words revealed higher phase similarity for remembered items at 5 Hz (cluster-randomization: $P = 0.016$, corrected for multiple comparisons) over right temporal electrodes. To examine the time-course of the reactivation effect, similarity-measures were again averaged across significant electrodes and t-statistics were computed for every time-point. Two distinct reactivation episodes emerged, peaking at 430ms (t_{16}= 5.16, $P = 0.00009$) and 2000ms ($t_{16} = 4.62$, $P = 0.0002$).

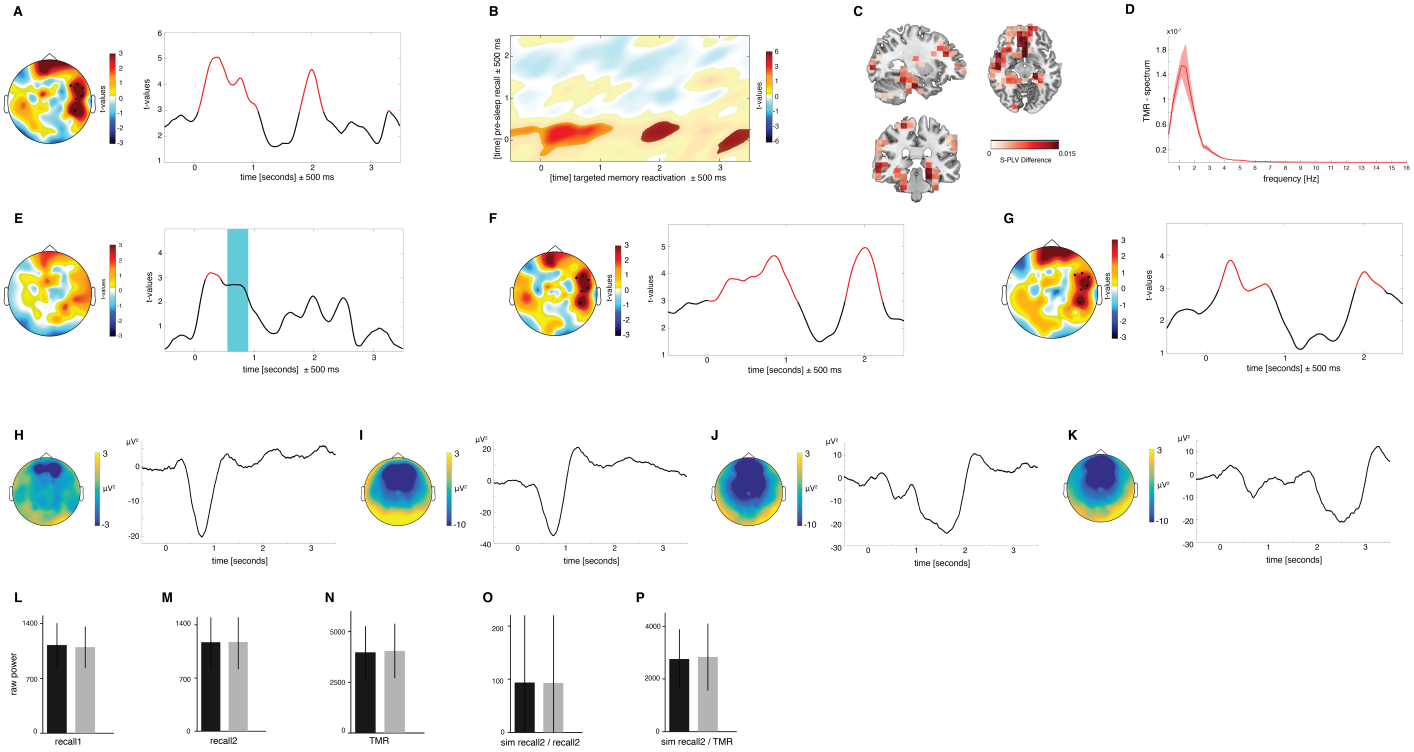


Figure S3 related to Figures 2+3. (A-E) Assessing word specific phase-similarity between recall2 and TMR up to 3.5 seconds. Due to the length of our TMR data segments and the 1 second width of the sliding time-window (from retrieval), robust similarity values could be obtained up to 2.5 seconds after cue presentation with regards to the TMR data. For any further time-point after 2.5 seconds a decreasing amount of data was available, thereby adding noise to the measurements. Still, to characterize the perpetuating effects of TMR induced reactivation-processes we investigated similarity measures up to 3.5 seconds (even though values after 2.5 seconds have to be taken with caution). **(A)** Re-current reactivation of recall-related phase-patterns at 5 Hz during TMR emerged over right temporal electrodes. The topography displays the test-statistics of the averaged difference in phase similarity between remembered and not-remembered words (0 - 3.5 seconds). The time-course depicts t-values averaged across highlighted electrodes (n = 6). The phase similarity at a given time point reflects the similarity computed in a window of +/- 500ms around this time point. **(B)** Assessing phase-similarity at 5 Hz between every time-point of retrieval and TMR confirmed the reoccurring pattern of similarity. Three distinct reactivation episodes emerged, peaking at 390ms (t_{16}= 4.49, P = 0.0003), 1990ms (t_{16} = 4.59, P = 0.0002) and 3310ms (t_{16} = 3.31, P = 0.004). **(C)** Source reconstruction (DICS beamformer) of the effects shown in (a). The difference in phase similarity for remembered and not-remembered items indicates large differences in right (para)hippocampal regions and left-frontal areas. **(D)** Frequency spectrum of the TMR similarity measures showed a ~1 Hz periodicity of reactivation processes. Shading denotes SEM **(E)** In line with behavioral predictions, providing a target stimulus after the TMR-cue blocked associated reactivation processes. The time course depicts t-values averaged across highlighted electrodes in (a). Presentation of the target word is highlighted in petrol. Only a brief reactivation effect at 270ms (before target word onset) emerged. The topography displays the test-statistics of the averaged difference in phase similarity between remembered and not-remembered words (0 - 3.5 seconds). No significant cluster was found. **(F+G) Varying the sliding time window width. (F)** Using a sliding time-window containing the temporal pattern from 'recall 2' with a width of 800 ms (4 cycles) to obtain phase similarity between recall and TMR indicates a highly comparable results pattern as obtained in the main analysis, with a re-current reactivation of recall-related phase-patterns at 5 Hz during TMR emerging over right temporal electrodes. The topography exhibits the test-statistics of the averaged difference in phase similarity between remembered and not-remembered words (0 - 2.5 seconds). To examine the time-course, similarity-measures were averaged across significant electrodes (n = 6) and t-statistics were computed for every time-point. Two reactivation episodes emerged, peaking at 850 ms (t_{16}= 4.66, P = 0.0002) and 2010 ms (t_{16}= 4.96, P = 0.0001). **(G)** Using a sliding time-window with a width of 1200 ms (6 cycles) to obtain phase similarity between recall and TMR also indicates a re-current reactivation of recall-related phase-patterns at 5 Hz during TMR emerging over right temporal electrodes (the reduced robustness of the effects might indicate that the usage of a rather long sliding time-window comes at cost of specificity). The topography exhibits the test-statistics of the averaged difference in phase similarity between remembered and not-remembered words (0 – 2.5 seconds). To examine the time-course, similarity-measures were averaged across significant electrodes (n = 6) and t-statistics were computed for every time-point. Two reactivation episodes emerged, peaking at 740 ms (t_{16}= 3.13, P = 0.0064) and 2010 ms (t_{16}= 3.5, P = 0.0029). **(H-K) Slow oscillation detection. (H)** Mean EEG signal of all detected SOs (at Fz) averaged (across subjects) time locked to cue presentation (t = 0 s). **(I)** Mean EEG signal of detected SOs during the first second (n = 168.29 ± 51.58) **(j)** second (n = 118.35 ± 39.49) and **(K)** third second (n = 118.35 ± 39.49 after stimulus onset (t = 0 s). A repeated measures ANOVA indicated a significant difference between the number of detected SOs for the three time segments ($F_{(2,32)}$= 31.26, P < 0.001). Post-hoc t-tests revealed that the number of SOs occurring in the first second was significantly higher as compared to the second time-segment (t = 5.31, P < 0.001), while the same direction was evident when comparing the second and third second (t = 3.85, P < 0.001). **(L-P) Spectral power control analyses.** There was no significant power-difference between remembered and non-remembered items with regards to **(L)** recall1 (P > 0.14) **(M)** recall2 (P > 0.15) and **(N)** TMR (P > 0.35; all results were controlled for multiple comparisons across space and time using cluster-based permutation statistics). We further estimated oscillatory power for the very same contrasts as utilized in the phase similarity analyses for recall1 vs. recall2 **(O)** and recall2 vs. TMR **(P)**, respectively (see Supplementary Methods for details). No significant differences were observable in both analyses (P >0.8; P >0.35; results were controlled for multiple comparisons across space and time using cluster-based permutation statistics). For illustrative purposes oscillatory power was averaged over the electrodes of interest. Note that increasing the frequency range to 3-7 Hz also yielded no significant differences for all the analyses displayed here (all p's > 0.3).

Supplemental Experimental Procedures:

Task and Procedure

German speaking participants, performed a vocabulary-learning task in the evening (~10pm). The task consisted of 120 Dutch words and their German translation. With regards to the first learning round, each trial consisted of a Dutch word, which was succeeded by the German translation. All words were presented via loudspeaker. The trials of the second round (referred to as 'recall1') started with the presentation of the Dutch word (cue), followed by a question mark for up to 7 seconds. The participants were asked to vocalize the correct German translation if possible (target). within the 7 seconds (if possible) or to indicate if they were not able to do so. In any case, the correct German translation was presented afterwards. The same cued recall procedure was accomplished in the third round ('recall2'), except that performance feedback was omitted. Participants correctly recalled 43.47 ± 2.56 out of 120 words during recall1 and 62.05 ± 3.22 (out of 120) during recall2. The learning phase was followed by a 3 hours retention interval of sleep.

During NREM sleep, subsets of the Dutch cue words learned before the retention interval were repeatedly presented for 90 minutes via loudspeaker. Cues were presented every ~4 seconds during stable NREM sleep (sleep stages N2 and SWS) via loudspeaker (50 dB sound pressure level) summing up to ~ 16 repetitions per word in random order. To assure that words were exclusively presented during NREM sleep (sleep stages N2 and SWS), sleep was permanently monitored by the experimenter. The stimulation protocol was manually interrupted whenever signs of arousals, awakenings or REM sleep were visible. Memory cues were presented either as single cues (only the Dutch words), 40 as word pair cues (Dutch and German words) and 40 were not replayed at all. During word pair cueing, the presented word-pairs consisted of correct word-pairs just as learned before the retention interval for 8 participants, while 9 participants received incorrect targets following the cues (as compared to learning). The false cue-target combinations were created by randomly intermixing the Dutch and German words of this category. Thus, new Dutch–German word combinations for presentation during sleep were formed.

Importantly opposed to replaying single cues, cueing of word pairs, irrespective of the category (correct, wrong), was associated with a suppression of the beneficial effects of cueing. Thus, we capitalized our main analysis on the single-cue TMR condition, while the word-pair condition served as control. In all three categories, the relation of remembered and non-remembered word pairs of the last learning trial before sleep was maintained. Hence, all categories comprised the same number of remembered and non-remembered words before sleep. All words were individually and randomly chosen for each participant using an automatic MATLAB algorithm. After sleep, recall of the vocabulary was tested in a final retrieval phase using a cued recall procedure. For the details on EEG recording and pre-processing see Supplementary Methods. All of the following steps were accomplished with MATLAB (the MathWorks) using the open-source FieldTrip toolbox (Oostenveld et al., 2011).

Supplemental Methods:

EEG recording and preprocessing

EEG was recorded using a high-density 128-channel Geodesic Sensor Net (Electrical Geodesics, Eugene, OR). Impedances were kept below 50 kΩ. Voltage was sampled at 500 Hz and initially referenced to electrode Cz. Offline EEG preprocessing was realized using BrainVision Analyzer software. Data were re-referenced to an average-reference. The continuous EEG was epoched into intervals from 1,000ms before until 3,000ms after word onset. Trials affected by muscle or movement artifacts were manually

removed. Eye blinks and movements of the pre-sleep EEG recordings were corrected using independent component analysis (Jung et al., 1998). For each phase (recall1, recall2 and TMR), segments were categorized based on the subjects' memory performance in the final retrieval phase into later remembered and non-remembered words..

Single-trial Phase Locking Value

The similarity of two signals was assessed as 1 minus the circular variance of difference in phase over time (Lachaux et al., 2000). The S-PLV is robust with regards to noisy data and allows for assessing similarity between two time windows in non-time-locked data (Michelmann et al., 2016). S-PLV ranges from 0 to 1, with 1 indicating perfect phase locking. Phase values were extracted using a complex Morlet wavelet of 6 cycles for frequencies between 1 and 20 Hz in steps of 0.5 Hz between 1000ms pre- to 3000ms after stimulus onset. For computational efficiency, the resulting phase values were down-sampled to 100 Hz. Lachaux and colleagues (2000) recommend assessing the S-PLV over 6–10 cycles of a given frequency to receive a good signal-to-noise ratio. At 5 Hz this would have resulted in a width of 1200 – 2000ms. Given that content-specific phase similarity during recall was observable in a rather narrow time window, we decided to use a 1 second sliding time-window (including 5 cycles). This should allow for obtaining an acceptable signal-to-noise ratio, while maintaining sufficient specificity (for the effects of varying the widths see Figure S3 (f + g)).

Slow Oscillations Detection

SOs were detected in all TMR data segments in EEG channel Fz. The signal was band-pass filtered between 0.5 and 4.0 Hz. The detection procedure resembled those earlier described in Bölsterli et al (2011) and Riedner et al., (2007). Half-waves were defined as negative deflections between two zero crossings. Corresponding quarter-waves were defined as the time between the negative peak and the following zero-crossing. Negative half-waves reaching at least an amplitude of −75μV were included in the analysis. Furthermore, only half-waves with a duration between 0.25 and 1 second (0.5–2 Hz) and which had a corresponding quarter-wave with a duration of more than 0.11 s (<2.25 Hz) were considered as SOs.

Source estimation

A spatial filter for each specified location (each grid point; $10mm^3$ grid) was computed based on the cross -spectral density, calculated for 5 Hz, using a complex Morlet wavelet, for all trials. As time windows of interest served the 1 second time-window derived from 'recall2' and the first second of TMR, given the most extended pattern of phase similarity effects during this period. Electrode locations for the 128-channel Geodesic Sensor Net EEG system were co-registered to the to the surface of a standard MRI template in MNI (Montreal Neurological Institute) space using the nasion and the left and right preauricular as fiducial landmarks. A standard leadfield was computed using the standard boundary element model (Oostenveld et al., 2003). The forward model was created using a common dipole grid ($10mm^3$ grid) of the grey matter volume (derived from the anatomical automatic labeling atlas (Tzourio-Mazoyer et al., 2002) in MNI space, warped onto standard MRI template, leading to 1457 virtual sensors. Data analysis was accomplished in the same way as before on sensor level.

TMR spectrum

As TMR during sleep seemed to have triggered reactivation processes in a recurrent fashion, we evaluated whether the similarity measures would fluctuate at a certain frequency .We performed a spectral analysis of the time-course of the phase similarity difference. We estimated the spectral power

for frequencies between 0.25 and 16 Hz by multiplying a hanning taper to the Fourier transformation (-0.5 to 2.5 seconds) and evaluated potential peak frequencies, after correcting for the 1/f shape of the power spectrum. To correct for the 1/f shape of the power spectrum, this was accomplished on the first time-domain derivative of the data. Thereby power at every frequency bin is multiplied by its frequency, neutralizing the 1/f effect (Sleigh et al., 2001).

Cluster-based nonparametric permutation tests

This approach controls the Type I error rate with regard to multiple comparisons by clustering neighboring sensor pairs, exceeding a critical t-value in the same direction. For all included frequency bins (3-16 Hz), paired sampled t-tests were computed for any given electrode and for each time-point (-0.5 to 2.5 seconds). Thereby, clusters of contiguous sensors across participants were identified ($P < 0.05$, two-tailed). The cluster-level statistic was defined from the sum of the t values of the sensors in a given cluster. Only the cluster with the largest summed value was considered and tested against the permutation distribution (Monte-Carlo method, $P < 0.05$, two-tailed t-test).

Spectral power control analyses

As power differences can bias phase estimation, we tested whether there was a significant difference in power between conditions. We evaluated potential differences in oscillatory power for remembered versus non-remembered items, with regards to all steps of the experiment. Power values were extracted for 5 Hz using a complex Morlet wavelet of 6 cycles with regards to 2.5 seconds following stimulus onset. Oscillatory power for words remembered after 'recall2' was subtracted from power values of 'recall1' and contrasted against the difference of power-values of words not remembered during 'recall2' and their equivalent during 'recall1'. We specifically tested potential differences in oscillatory power for electrodes, which showed significant effects in phase-similarity between wakefulness and sleep (see Fig.3a; but please not that similar outcomes were obtained when adding the overall number of electrodes). We further estimated oscillatory power for the very same contrasts as utilized in the phase similarity analyses for recall1 vs. recall2 and recall2 vs. TMR respectively. With regards to the latter analysis oscillatory power for words remembered after sleep was subtracted from power values of 'recall2' and contrasted against the difference of power-values of words not remembered after sleep and their equivalent during 'recall2'.

References

Jung, T.-P., Humphries, C., Lee, T.-W., Makeig, S., McKeown, M.J., Iragui, V., and Sejnowski, T.J. (1998). Extended ICA removes artifacts from electroencephalographic recordings. In Advances in Neural Information Processing Systems, pp. 894–900.

Lachaux, J.-P., Rodriguez, E., Le van Quejen, M., Lutz, A., Martinerie, J., and Varela, F.J. (2000). Studying Single-Trials of Phase Synchronous Activity in the Brain. Int. J. Bifurc. Chaos *10*, 2429–2439.

Michelmann, S., Bowman, H., and Hanslmayr, S. (2016). The Temporal Signature of Memories: Identification of a General Mechanism for Dynamic Memory Replay in Humans. PLoS Biol. *14*, e1002528.

Oostenveld, R., Stegeman, D.F., Praamstra, P., and van Oosterom, A. (2003). Brain symmetry and topographic analysis of lateralized event-related potentials. Clin. Neurophysiol. 114, 1194–1202.

Sleigh, J.W., Steyn-Ross, D.A., Steyn-Ross, M.L., Williams, M.L., and Smith, P. (2001). Comparison of changes in electroencephalographic measures during induction of general anaesthesia: influence of the gamma frequency band and electromyogram signal. Br. J. Anaesth. 86, 50–58.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., and Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15, 273–289.