# Difficulty in artificial word learning impacts targeted memory reactivation and its underlying neural signatures

Klaassen, A.-L.[1,2], & Rasch, B.[1]

1) Department of Psychology, Division of Cognitive Biopsychology and Methods, University of Fribourg, Fribourg, Switzerland
2) Department of Anesthesiology & Pain Medicine, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland

Corresponding author:
Björn Rasch
University of Fribourg, Department of Psychology, Division of Biopsychology and Methods, Rue P.A. de Faucigny 2, 1701 Fribourg, Switzerland.
bjoern.rasch@unifr.ch

**Abstract**

Sleep associated memory consolidation and reactivation play an important role in language acquisition and learning of new words. However, it is unclear to what extent properties of word learning difficulty impact sleep associated memory reactivation. To address this gap, we investigated in twenty-two young healthy adults the effectiveness of auditory targeted memory reactivation (TMR) during non-rapid eye movement sleep of artificial words with easy and difficult to learn phonotactical properties. Here, we found that TMR of the easy words improved their overnight memory performance, whereas TMR of the difficult words had no effect. By comparing EEG activities after TMR presentations, we found an increase in slow wave density independent of word difficulty, whereas the spindle-band power nested during the slow wave up-states - as an assumed underlying activity of memory reactivation - was significantly higher in the easy/effective compared to the difficult/ineffective condition. Our findings indicate that word learning difficulty by phonotactics impacts the effectiveness of TMR and further emphasize the critical role of prior encoding depth in sleep associated memory reactivation.

**Keywords:** targeted memory reactivation, prior knowledge, sleep-dependent memory consolidation, phonotactics, sleep spindles, alpha oscillations

## Introduction

Learning of new words is essential in language acquisition. During the learning process of unknown words, top-down influences can facilitate perception and encoding of these words depending on the similarity between the new words and the pre-existing knowledge. One critical characteristic of similarity to pre-existing knowledge in auditory word processing is its speech sound (phoneme) pattern. In phonology as the field of language specific phoneme structures, phonotactics determines the constraints of word phoneme composition of a specific language. Previous studies have shown that artificial words with similar phonotactic properties to real words are easier to learn and to remember[1,2], suggesting that pre-existing knowledge about the sound structure (i.e. schemas) facilitates encoding and consolidation of new words. A synchronized interplay between the ventromedial prefrontal cortex, hippocampus and unimodal associative cortices is proposed to explain schema instantiation and schema mnemonic effects[3] at the neural level. To describe the influence of phonotactics on memory processing within this framework, new words with phoneme structures congruent to activated schema variables are inserted with less cognitive effort and difficulty into pre-existing knowledge networks of long-term memory in contrast to schema-incongruent words.

According to the complementary system account by Gaskell and colleagues[4,5], learning of new words comprises a first stage of rapid initial familiarization represented by neural activities of the medial temporal lobe and a second stage of slow offline consolidation of the neocortex. The model suggests that, after initial hippocampal learning, through sleep consolidation memory representations, the words are gradually integrated into long-term lexical and phonological memory structures of the neocortex. Additionally, these processes of memory consolidation might depend on spontaneous and repeated reactivation of the newly formed memory traces during non-rapid eye-movement (NREM) sleep[6]. The active system consolidation hypothesis proposes synchronized hippocampal sharp-wave-ripples, thalamocortical sleep spindles and slow oscillations of NREM sleep as underlying activities of memory reactivation. Accordingly, studies in humans[7–9] and in rodent models[10,11] found that sleep spindles and hippocampal sharp wave ripples occur grouped during slow oscillations and that these events correlate with memory consolidation[6,12–15].

While similarity of the learning material to pre-existing knowledge clearly facilitates encoding, this notion is controversial for the effects on sleep consolidation. For example, Havas and co-authors[16] tested learning and consolidation of artificial words phonologically derived either from participant's native language (L1 words in Spanish) or foreign language (L2 words in Hungarian). Here, sleep-associated improvements in memory recognition were only observable in the condition of the L2 words, suggesting that a reduced similarity to pre-existing knowledge facilitates consolidation during sleep. Also, in a study by Payne et al.[17], sleep-associated memory benefits occurred only for semantically unrelated word-pairs, but not for semantically related words.

On the other hand, Zion and colleagues[18] reported that sleep facilitated consolidation in a second language learning task in participants with a higher degree of meta-linguistic knowledge, indicating a beneficial effect of pre-existing knowledge on new word consolidation. Similarly, Durrant and colleagues[19] reported that only schema-conformant but not non-conformant learnings benefitted from sleep. In a study from our own lab, German-participants learning unknown Dutch vocabulary profited more from sleep compared to native French

speaker when given the same amount of learning trials, whereas German native speakers might have more pre-existing knowledge about the Dutch words in relation to the French native speakers, which facilitated learning and consolidation of the new Dutch vocabulary[20].

A promising approach to examine sleep-associated memory consolidation is to experimentally bias reactivation by re-presenting olfactory[21–23] or auditory[24,25] reminder cues during sleep, a technique known as targeted memory reactivation (TMR). TMR has been related to elevated slow-wave and spindle activity following cueing presentations as neural markers of memory reactivation[26,27]. So far two studies suggest that prior knowledge increase the effectiveness of TMR: Groch and colleagues[28] reported improved memory performance by TMR for associations between pseudo-words and familiar objects, but not for word associations of new/unknown objects. Here, before sleep, subjects showed as well better encoding performance for the familiar compared to the unknown objects. In another study[29] of learning sound-object locations, the effectiveness of TMR correlated positively with the initial encoding performance prior to sleep. Additionally, both studies[28,29] found significant associations between TMR's effectiveness and sleep spindle activities after the cueing presentations. These findings suggest that higher similarity between the learning material and the pre-existing knowledge facilitates reactivation and consolidation during sleep. However, it is unknown whether these findings can be generalized to learning of new artificial words with different levels of phonotactical similarity to prior word-sound knowledge, as a factor of learning difficulty.

To address this issue, we designed artificial words with varying difficulty levels by phonotactical properties. To consider the beneficial effect of reward related information on sleep dependent memory consolidation and reactivation[30–33], we trained healthy young participants to categorize these words into rewarded and unrewarded words to gain and to avoid losses of money points. By using auditory TMR, we re-presented the artificial words during subsequent sleep to study its effect on categorization performance. As a manipulation check, we expect that artificial words with more familiar phonotactical properties would be easier to learn compared to words with less familiarity. Regarding previous findings[28,29], we hypothesized higher effectiveness of TMR for the easy to learn words in comparison to the difficult words, accompanied by higher oscillatory activity on the slow-wave and spindle range.

## Results

### Experimental design

To study the impact of difficulty in word learning on TMR, we developed a novel learning paradigm. We formed four sets of artificial words (40 words per set; see Table S3 and S4) consisting of different sequences of two vowels and two consonants. Here, we subdivided the alphabet into two groups of consonants (C1: b, c, d, f, g, h, j, k, l, m; C2: n, p, q, r, s, t, v, w, x, z) and vowels (V1: a, e, I; V2: o, u, y). Four-letter-words were created by selecting letters from the vowel and consonant groups according to four different sequences (G1:C1, V1, V2, C2; G2: C1, V1, C2, V2; G3: V1, C1, C2, V2; G4: V1, C1, V2, C2; Fig. 1a; see methods for further details). Comparison analyses between the sets revealed significant differences in phonotactic probability (PP; Fig. 1b; unpaired *t*-tests: G1 / G2 > G3 / G4, $p < 0.005$, values of Cohen's $d > 0.71$). PP quantifies the frequency of a single phoneme (phoneme probabilities) or a sequence of phonemes (e.g., biphone probabilities) within a language and thus serves as a measurement of the similarity between the artificial words and the pre-existing real word knowledge.

According to distinct levels of PP, we paired the four sets to the high- (G1 & G2) and low-PP (G3 & G4) condition, respectively.

During encoding, we trained the subjects to discriminate the artificial words based on reward associations by manual button presses. Therefore, the high- and the low-PP condition, consisted of 40 rewarded and 40 unrewarded words. As a two alternative forced-choice task, we assigned left- and right-hand button presses to the rewarded and the unrewarded word category, counterbalanced across subjects. We instructed the participants to respond to each word by left- or right-hand button presses, whereas one button means the word is rewarded (gain of money points) and the other button means the word is unrewarded (avoid the loss of money points). In the beginning, they had to guess. By three presentations of each word in randomized order and by feedback at the end of each trial, they learned to respond correctly according to the rewarded/unrewarded associations of the words (Fig. 1c).

To obtain a measurement of discrimination memory with respect to the potential influence of the response bias, we applied the signal detection theory[34]. Because, we instructed the participants to respond to each word by left- or right-hand button presses and that one button means reward is present whereas the other button means reward is absent, we considered correct responses of words of the rewarded category as hits and words of the unrewarded category as correct rejections. Accordingly, we assigned the responses with regard to the reward associations of the words to the following four response types: hits (rewarded, correct); correct rejections (unrewarded, correct); misses (rewarded, incorrect); and false alarms (unrewarded, incorrect). Dependent on responses, subjects received money points (Fig. 1d).

After encoding and before sleep, we tested the pre-sleep memory performance (Fig. 1e). Subjects slept one night in the sleep laboratory and during the NREM sleep stages 2 & 3 we conducted auditory TMR of the low-PP words in one group of subjects and TMR of the high-PP words in the other group (between subject design). At the following morning, we tested the post-sleep memory performance.

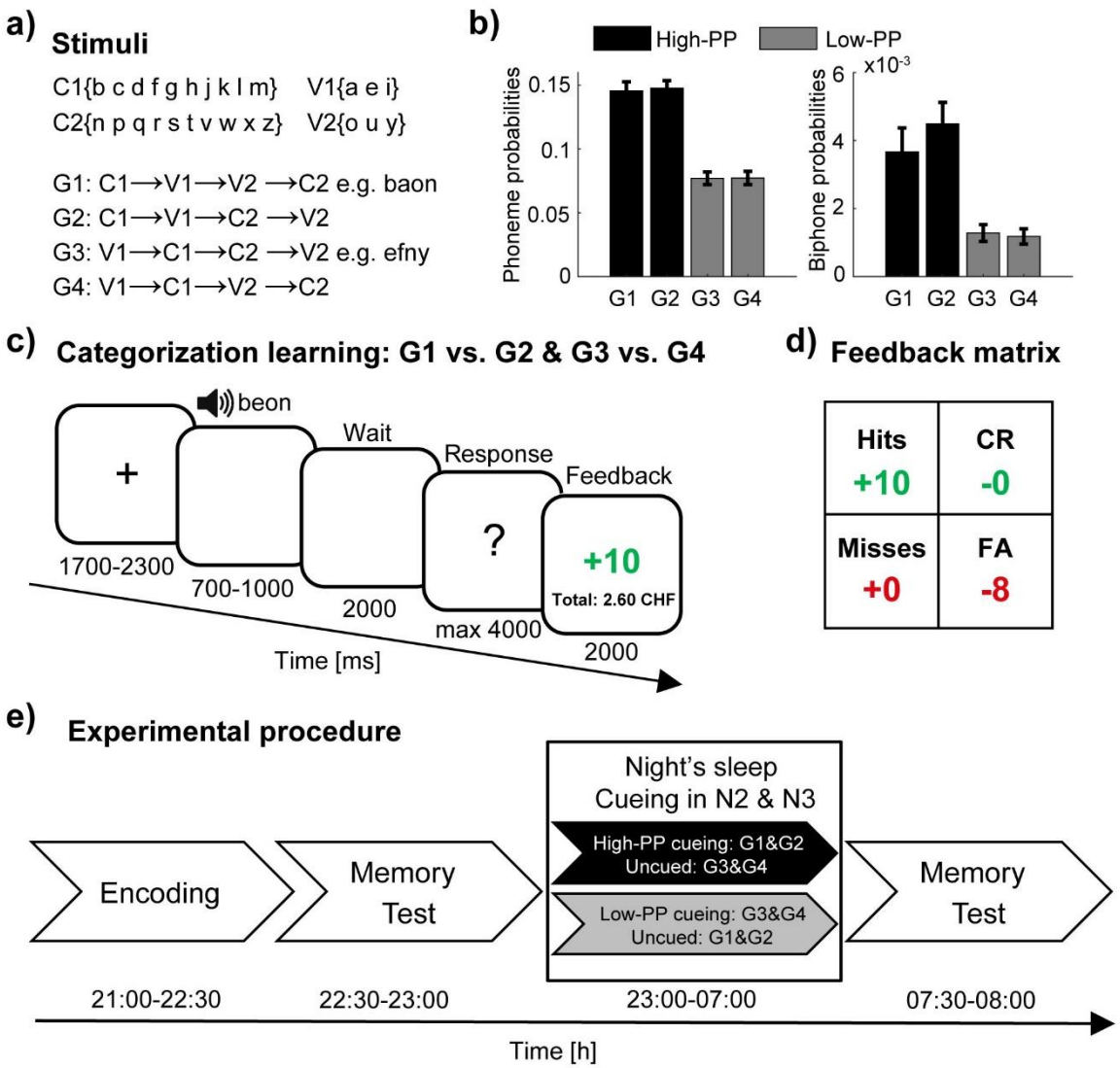



**Fig. 1. Experimental design.** a) Artificial words consist of two vowels and two consonants according to four different sequences. b) The phonotactic probabilities (PP) for single phonemes (left panel) and biphone probabilities (right panel) averaged values of the four sets (40 words per set) and pairing of the sets with respect to two distinct levels of PP in high- (black) and low-PP (gray). c) Schematic trial structure of the learning task with screen images and the durations in milliseconds. As a two alternative forced-choice task, responses of left- and right-hand button presses were assigned to the rewarded and the unrewarded word category, respectively. The participants were instructed to respond to each word by left- or right-hand button presses, whereas one button means the word is rewarded (gain of money points) and the other button means the word is unrewarded (avoid the loss of money points). d) Feedback matrix with the four answer types (hits: rewarded and correct; CR, correct rejections: unrewarded and correct; misses: rewarded and incorrect; FA, false alarms: unrewarded and incorrect) regarding to response and reward assignment of the word. Note, subjects could receive and lose money points dependent on correct and incorrect responses. e) Experimental procedure with experimental tasks and phases in temporal order. TMR took place in the NREM sleep stages 2 & 3. Error bars reflect standard errors of the mean (SEM).

## Manipulation check based on encoding analyses

To validate our novel paradigm, we examined the influence of PP on encoding performance. Based on the signal detection theory[34], we calculated d'-values to measure subject's abilities to differentiate and categorize rewarded and unrewarded words of the two PP conditions over the three presentations of the encoding task (Fig. 2a). A repeated-measure ANOVA on d' with PP

(high vs. low) and presentations (1 to 3) as within-subjects factors revealed significant main effects of PP ($F(1,32) = 5.13$, $p = 0.03$, $\eta^2 = 0.14$), and presentations ($F(2,64) = 95.67$, $p < 0.001$, $\eta^2 = 0.75$). Additional pairwise comparisons between PP conditions showed significant differences after the first presentation (paired *t*-tests of the three presentations in order: (1) $t(32) = 0.73$, $p = 0.47$; (2) $t(32) = 2.35$, $p = 0.03$, Cohen's $d = 0.41$; (3) $t(32) = 2.17$, $p = 0.04$, Cohen's $d = 0.38$). These results indicate superior learning performance for words with high- compared to words with low-PP.

To examine the oscillatory correlates of encoding performance, we conducted time-frequency analyses time-locked to auditory word representations (Fig. 2b) and extracted power values over time, frequency, and EEG electrodes by using wavelet transformations (see methods). This analysis revealed a strong desynchronization after word presentations (0.7 to 1.9s to stimulus onset) in the frequency range of the alpha waves (8-13 Hz) and pronounced over parietal and occipital electrodes in contrast to baseline activity (-1 to -0.1s to stimulus onset; Fig. 2c). Testing whether changes in alpha desynchronization correspond to increased encoding performance over repetitive presentations, a repeated-measure ANOVA on alpha power values merged over central electrodes (Pz, Cz) including presentations (1 to 3) as a single within-subjects factor revealed a significant increase in alpha desynchronization ($F(2,64) = 4.19$, $p = 0.02$, $\eta^2 = 0.12$).

Regarding distinct levels of encoding performance between high- and low-PP, we compared next changes in alpha desynchronization between these conditions. We obtained the changes in power values by subtracting the first from the second and third presentation for the high- and low-PP condition, respectively. Here, the first word presentation of naive stimulus processing served us with a more representative baseline condition covering the time-window of interest of 0.7 to 1.9 s after the stimulus onset to examine relevant changes of encoding. Cluster-based statistics of the third presentation revealed a significant cluster over the left posterior electrodes with a stronger alpha power decrease for the high- in contrast to the low-PP condition (time window: 0.7 to 1.9s; frequency range 8-13Hz; averaged *t*-value over the 5 cluster electrodes = -2.69, $p = 0.01$, Cohen's $d = -0.48$; Fig. 2d). A post-hoc repeated-measure ANOVA on alpha power changes (merged over Pz and Cz electrodes) with PP (high vs. low) and presentations (2 to 3) as within-subjects factors revealed a main effect of PP ($F(1,32) = 5.42$, $p = 0.03$, $\eta^2 = 0.15$), and a significant interaction ($F(1,32) = 7.38$, $p = 0.01$, $\eta^2 = 0.19$; Fig. 2e). Additional post-hoc pairwise comparison between PP conditions showed significant difference of the third presentation (paired *t*-tests of presentations in order: (2) $t(32) = -0.36$, $p = 0.72$; (3) $t(32) = -3.55$, $p = 0.001$, Cohen's $d = -0.62$) and a *t*-test against 0 revealed significantly decreased activity in high-PP of the third presentation ($t(32) = -2.92$, $p = 0.006$, Cohen's $d = -0.51$). In addition to the behavior results, these EEG results indicate differences between PP conditions in desynchronization of alpha oscillations, as an assumed neural correlate of encoding depth[35,36]. To summarize, as a manipulation check based on encoding analyses, we confirmed that the conditions of high- and low-PP correspond to distinct levels in learning difficulty.

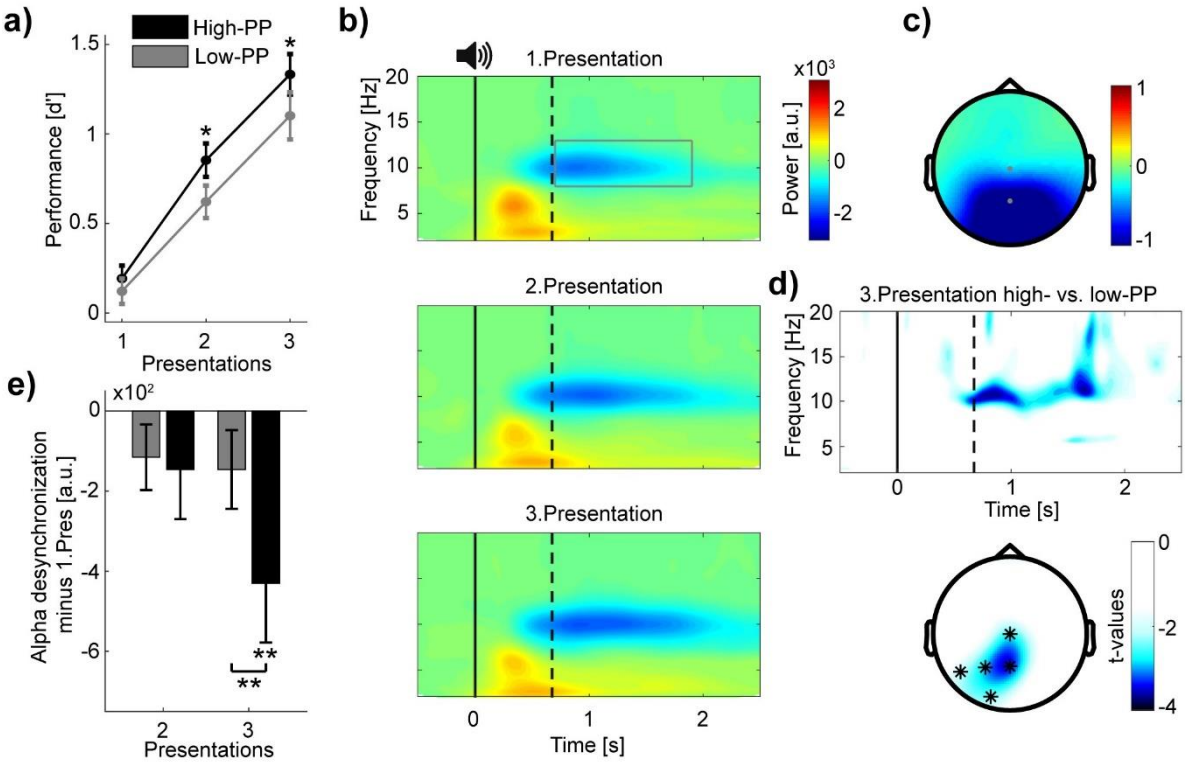



**Fig. 2. Distinct levels of encoding between high- and low-PP words.** a) Learning curves showing encoding performance over presentations of high- (black) and low-PP (gray) words. Note significant greater performance of high- in comparison to low-PP over the second and third presentations. b) Grand average time-frequency plots time-locked to word presentations. The gray rectangle within the top panel borders time and frequency range of interest (0.7 to 1.9s; 8-13Hz). Three different panels from top to down regarding to the three presentations. Solid and dotted lines within plots representing stimulus onset and averaged offset, respectively. Note, increases of oscillatory desynchronization in the alpha range (8-13Hz) over the three presentations. c) Topographic map shows power values averaged over the time window and frequency range of interest. d) (Top) Time-frequency representation of *t*-values (merged over Pz and P3 electrodes) shows significant greater changed desynchronization in alpha oscillations of high- in contrast to low-PP during the third presentation. Below, topographic map indicates significant cluster of electrodes of comparison between PP conditions of the third presentation (0.7 to 1.9s; 8-13Hz). e) The bar chart shows mean changes across subjects in alpha power (merged over Pz and Cz electrodes) of the second and third presentation by subtracting the first presentation in high- (black) vs. low-PP (gray). Statistical analyses revealed significant higher desynchronization of high- compared to low-PP and a significant decrease in alpha power under 0 of high-PP at the third presentation. Error bars reflect standard errors of the Mean (SEM); *$p$<0.05, **$p$<0.01.

## TMR affects memory consolidation of the easy to learn words

To examine whether TMR during sleep impacts memory consolidation of discrimination learning with respect to learning difficulty, we calculated the overnight changes by subtracting the pre- from the post-sleep memory performance based on d'-values of the reactivated sequences (cued) and non-reactivated sequences (uncued). In group 1, the high-PP sequence was cued, and the low-PP sequence was uncued. In group 2, the low-PP sequence was cued, and the high PP sequence was uncued (see Fig. S5 for pre- and post-sleep memory test data of the different conditions). Conducting *t*-tests against 0 revealed a significant increase in the TMR/cueing condition of group 1 (high-PP cued: 0.24 ± 0.1, $t(10) = 2.44$, $p = 0.035$, Cohen's $d = 0.74$), while the memory performance remained unchanged in all the other conditions (low-PP cued: -0.09 ± 0.08, $t(10) = -1.16$, $p = 0.27$; high-PP uncued: 0 ± 0.08, $t(10) = -0.04$, $p = 0.97$; low-PP uncued: 0.03 ± 0.1, $t(10) = 0.31$, $p = 0.76$; Fig. 3a).

An additional two-way mixed design ANOVA on the same values with the factor cueing (cued vs. uncued) as a within-subject factor and group as a between-subject factor revealed trends of significance ($p < 0.1$) for the interaction (cueing × group: $F(1,20) = 3.47$, $p = 0.08$) and the main effect of group ($F(1,20) = 3.28$, $p = 0.09$). The main effect of cueing was not significant ($F(1,20) = 0.58$, $p = 0.46$). Post-hoc pairwise comparisons revealed a significant difference between the conditions of high-PP-cued and low-PP-cued ($t(20) = 2.63$, $p = 0.02$, Cohen's $d = 1.12$), a trend of significance between high-PP-cued and high-PP-uncued ($t(20) = 1.91$, $p = 0.07$) and no significant difference between high-PP-cued and low-PP-uncued ($t(10) = 1.55$, $p = 0.15$). In additional control analyses, pre-sleep memory performance and vigilance shortly before the post-sleep memory task did not significantly ($p$-values $\geq 0.08$) differ between both cueing groups (see Table S1).

We examined next as an exploratory analysis whether TMR conditions influence biases in decision-making. Therefore, we analyzed changes between the pre- and post-sleep memory tasks of c-criterion values as a measurement of risk avoidance/seeking in decision making[34]. Here, we found as well a significant overnight increase for cued sequences in group 1 (high-PP cued: 0.2 ± 0.07, $t(10) = 3.04$, $p = 0.01$, Cohen's $d = 0.92$), indicating an increased bias towards risk avoidance by TMR (Fig. 3b), while all the other conditions remained unchanged (high-PP uncued: 0.09 ± 0.06, $t(10) = 1.4$, $p = 0.19$; low-PP cued: 0.08 ± 0.05, $t(10) = 1.63$, $p = 0.13$; low-PP uncued: 0.03 ± 0.05, $t(10) = 0.56$, $p = 0.59$). An ANOVA on c-criterion changes showed no significant effects (interaction cueing × group: $F(1,20) = 2.66$, $p = 0.12$; main effect cueing $F(1,20) = 2.08$, $p = 0.17$; main effect group $F(1,20) = 0.38$, $p = 0.55$).

Taken together, these results suggest that the effectiveness of TMR depends on the level of difficulty in word learning, while auditory cueing during sleep increases the memory performance of the easy to learn words.

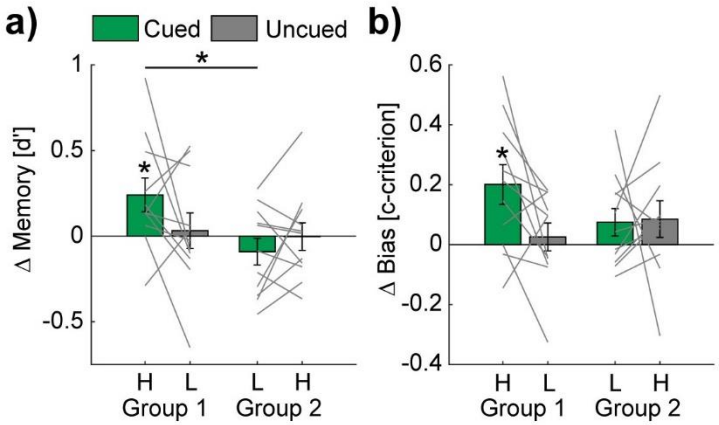



**Fig. 3. TMR affects the memory performance of the easy to learn words.** Bar charts show mean overnight changes of d' (a) and c-criterion (b) values of high (H) - and low (L) - PP and cued (green) vs. uncued (gray) conditions. Note, statistical analyses revealed significant overnight increases only in the high-PP cued condition. Gray lines represent individual data points. Error bars reflect SEM; *$p$<0.05.

## Increased spindle power nested during slow wave up-states in TMR of the easy to learn words

After analyzing TMR's effectiveness on behavior, we investigated the corresponding neural activities by EEG. By visual inspection of the signals, auditory word presentations during NREM sleep led to broad high-amplitude oscillations, called slow waves (SW; 0.5-3Hz),

whereas sleep spindle activity (9-16Hz) with various amplitude occurred preferentially nested during the SW's up-state phase (see for example traces Fig. 4a).

To statistically analyze whether TMR increases the SW density, we conducted a detection algorithm. To control for individual differences, we included 30% of SW with the highest amplitude per subject (see methods) in subsequent analyses. Testing changes of SW density in time windows of 0.5s, revealed significant increases after stimulus onset in comparison to baseline (Fig. 4b; baseline period of -1.5 to 0s before stimulus onset; *t*-tests against 0; 0 to 0.5s: 49.73 ± 14.32%, $t(21) = 3.47$, $p = 0.002$, Cohen's $d = 0.74$; 0.5 to 1s: 83.44 ± 14.97%, $t(21) = 5.57$, $p < 0.001$, Cohen's $d = 1.19$; 1.5 to 2s: 40.33 ± 10.67%, $t(21) = 3.78$, $p = 0.001$, Cohen's $d = 0.84$; see Table S6 for all statistical comparisons of time-bins from -0.5 to 3s and Fig. S7 for SW distribution from -0.5 to 6s time-locked to stimulus onset). Additional comparisons between the TMR conditions of high- and low-PP showed no significant differences in SW density, SW amplitude, number of TMR presentations and other sleep parameters, but a trend of significance for REM sleep parameters (see Table S2). In sum, these results indicate an increase of SW in response to auditory word stimulation independent of difficulty in word learning.

Next, we analyzed whether variations in spindle-band oscillations nested during the SW up-states after auditory TMR – as a neural signature of memory reactivation – differ between the conditions of high- and low-PP. To extract the spindle power during SW up-states, we conducted time-frequency analyses time-locked to the troughs of the detected SW in the time-window from 0 to 6s after auditory presentations. Fig. 4c shows the grand average analyses per condition, whereas the fast spindle band activity (12-16Hz) nested during SW up-states displayed a topography with maximum power over the central-parietal electrodes. Comparison between PP conditions revealed a significant greater spindle-band power prominent in the frequency range of 11-14Hz during SW up-states (from 0.3 to 0.8s after the SW trough) of high- in contrast to low-PP (averaged *t*-value over the 5 cluster electrodes = 2.66, $p = 0.01$, Cohen's $d = 1.13$; Fig. 4d), whereas post-hoc comparison including the individual SW amplitude as a covariate still revealed the significance of this cluster ($t(20) = 2.38$, $p = 0.03$, Cohen's $d = 1.02$).

By conducting an additional analyses based on detection of fast sleep spindles (12-16Hz; see methods), we confirmed that fast sleep spindles during the SW up-states (from 0.3 to 0.8s after the SW trough) occurred with significantly higher amplitude after the cueing presentation of high- compared to low-PP words, whereas parameters of sleep spindle density and the amount sleep spindles coupled to the SW up-state did not differed between the cueing conditions (see Fig. S8 and Table S9).

These results suggest that successful TMR of the easy to learn words is associated with increased spindle activity nested during SW up-states.

Whereas we found a significant group difference in spindle power nested during SW up-states, conducting further whole sample ($n = 22$) correlation analyses between the individual spindle power values of the significant cluster and the overnight changes of behavior measurements revealed no significant correlations (Δ d': $r = 0.16$, $p = 0.48$; Δ c-criterion: $r = 0.19$, $p = 0.40$).

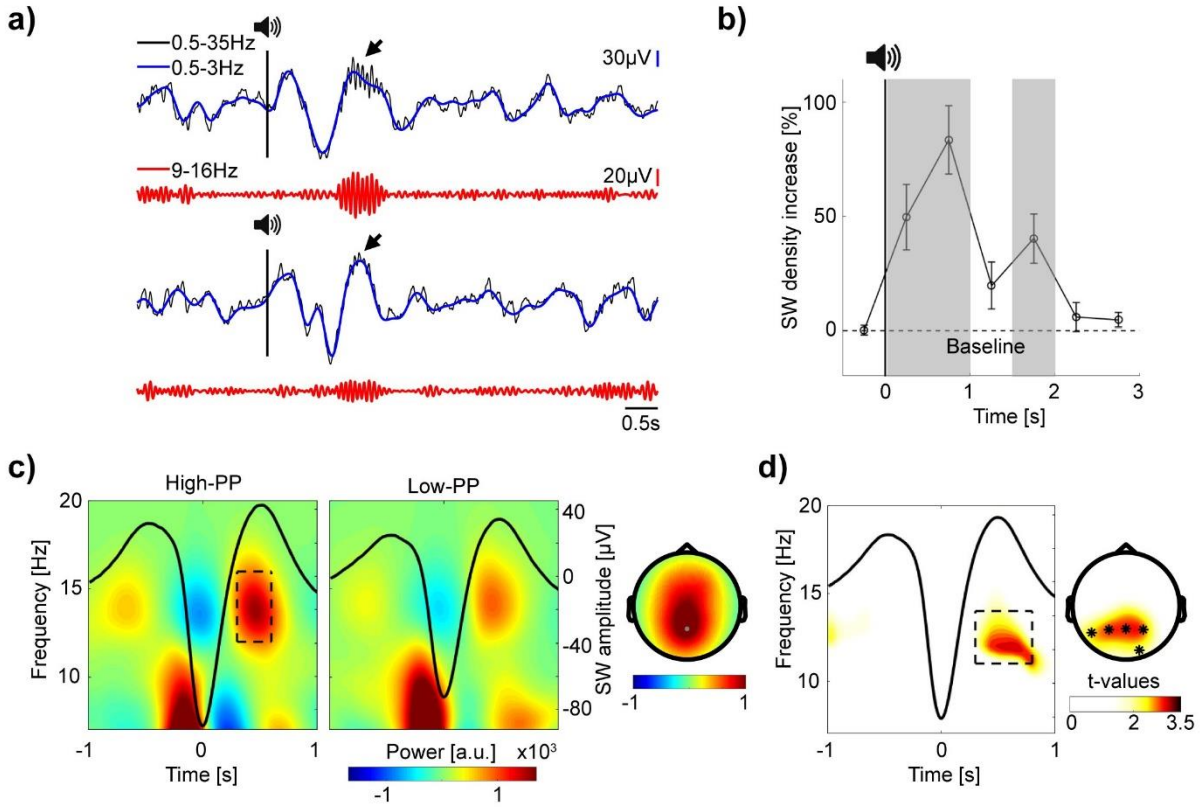



**Fig. 4. Increased spindle power during SW up-states in TMR of the easy to learn words**. a) Top and bottom panel, two example EEG traces of auditory cueing during sleep (-2 until 6s to stimulus onset). Top rows, in blue, signal filtered in the SW range (0.5-3Hz) superimposed upon the broadband (0.5-35Hz) signal in black. Vertical black lines with speaker symbols on top mark onsets of auditory presentations. Black arrows point to spindle activity during SW up-states. Bottom rows, in red, the same signal, but filtered in the spindle range (9-16Hz). Note, elevated SW following cueing presentations with various spindle band activity nested during SW up-states. b) Grand average baseline corrected curve of increased SW density after TMR in percentage. Shaded gray areas mark time windows (0 to 0.5s, 0.5 to 1s and 1.5 to 2s) of significant increased SW density. c) Grand average time-frequency plots time-locked to the troughs of SW with averaged signals plotted as black lines. Two different panels (left and right) according to high- vs. low-PP cueing conditions. The rectangle within the left panel borders time (0.3 until 0.6s to SW troughs) and frequency range of up-state fast spindle band activity (12-16Hz). Corresponding topographic map at right shows elevated fast spindle power over mid-parietal electrodes. d) Time-frequency representation of *t*-values time-locked to SW shows significant greater spindle band power during SW up-states for high- vs. low-PP (merged over Pz, P3 and P4 electrodes). Right, topographic map of *t*-values shows corresponding significant cluster of electrodes (0.3 to 0.8s; 11-14Hz), *$p$<0.05.

## Discussion

In this study, we asked whether learning difficulty of artificial words affects TMR's effectiveness and its neural activities. After confirming easier encoding of artificial words with high- compared to low phonotactic familiarity, we found that auditory TMR during sleep improved the memory performance of the easy to learn words, whereas cueing of the difficult words had no effect. Correspondingly, at the neural level, we observed that cueing presentations of the easy words induced significant greater spindle-band oscillations nested during SW up-states in comparison to the difficult words. To our knowledge, we present here the first study to address the critical role of learning difficulty by experimental manipulations of phonotactical word properties on TMR's effectiveness and its neural signatures.

Related to our finding of TMR's effectiveness of the easy words, previous research in object association learning showed beneficial influences of encoding depth[29] and prior knowledge[28] on TMR. Language studies without using TMR revealed that prior knowledge affected sleep dependent memory integration of new words in children[37] and in adults[18,20]. To explain these critical influences on consolidation during sleep, the model of the complementary systems approach of word learning (CLS)[4] proposes that new word memory temporary encoded in the hippocampus transfer during sleep into cortical structures of lexical and phonological memory for long-term embedding, due to co-activation of both memory systems. Thus, word memory reactivation and consolidation during sleep depend on the hippocampal encoding depth and the cortical prior knowledge in phonotactics. Regarding our results, we suggest facilitated systems co-activation/consolidation of the easy to learn words by auditory cueing during sleep as a manipulation of both factors - via enhanced initial encoding and higher similarity to pre-existing phonotactical knowledge - compared to the difficult to learn words.

Contrary to our results, other studies found better memorization of words with low phonological[16] and semantical[17] similarity to pre-existing knowledge after a retention period of sleep compared to words with high similarity. To explain this discrepancy, we could speculate that a very high degree of similarity led to a strong memory formation during wake-encoding, which prevented an additional effect of consolidation during sleep. Indeed, Creery et al.[29] showed that learned object-location associations only benefited from TMR during sleep when not almost perfectly recalled during prior wake. Here, a moderate encoding depth might benefit most from subsequent sleep consolidation compared to insufficient or almost perfect encoding. Alternatively, the contrary results could be related to characteristics of the learning paradigms (L1- vs. L2-like novel words with associations to objects[16]; semantically related vs. unrelated word pairs[17]) distinct to our design.

By conducting an additional exploratory analysis, we observed a significant change of the decision bias in the cueing condition of the easy to learn words and no overnight changes in the other conditions. Here, subjects became more biased in their decision making to avoid losses of money points. This observation suggests that TMR not only captures specific memory traces of the words and their reward associations, but also decision-making tendencies. In a previous study, Ai and colleagues[38] revealed altered reward-relevant decision preferences by TMR. To what extent concepts of decision-making such as risk avoidance/seeking could be reactivated, transformed, and consolidated by TMR is largely unexplored and might be interesting for future research for example in neuroeconomics.

Correspondent to our behavior results of TMR's effectiveness, we found significantly higher spindle-band oscillations nested during SW up-states after cueing presentations of the easy in comparison to the difficult words. Accordingly, the active system consolidation hypothesis proposes slow wave - sleep spindle coupling activities as an underlying mechanism of systems consolidation[6,12,39]. In addition to our result of the significant group difference, we failed to find significant correlations between SW nested spindle power values and overnight changes in behavior measurements, whereas previous studies reported associations of SW and spindle activities during sleep with the integration of new memories in pre-existing knowledge networks[40,41]. Strikingly, recent studies by implementing machine learning for EEG decoding revealed that classification accuracy of distinct memory categories peaked above chance level synchronized with SW-spindle coupling events[26,42]. Here, our results support the assumption of SW-spindle activities as an underlying mechanism of memory reactivation during sleep by

providing additional evidence to link successful TMR on behavior to increased spindle activity during SW up-states.

In validation of our novel paradigm, we confirmed easier encoding and better memorization of artificial words with high- in comparison to low-PP, in line with previous studies[1,2]. These results suggest that participants used their prior knowledge of frequently occurring word-sound patterns to facilitate learning. Regarding the framework of schema theory[3], processing of incoming words with high-PP might co-activate consistent schemata of prior word sound knowledge and thus may lead to facilitated memory integration and stabilization during encoding in comparison to processing of more inconsistent words with low-PP.

Our EEG analyses of encoding revealed desynchronization of alpha oscillations as a function of learning improvements. Additionally, superior encoding of the high-PP words corresponded to greater alpha desynchronization in comparison to the low-PP words. Consequently, we assume that the suppression of alpha oscillations over parietal and occipital electrodes represent a neural correlate of mnemonic functioning. Accordingly, previous studies linked desynchronization of alpha oscillations to memory processing[35,43] and memory replay during wake[44]. A study of intracranial electrophysiology in epilepsy patients[45] showed that neocortical alpha/beta desynchronization precede and predict fast hippocampal gamma activity during learning. The authors propose that these coupled activities represent hippocampal-cortical information transfer important for memory formation. Here, our findings let suggest cortical alpha desynchronization as neural marker of wake mnemonic processing according to distinct levels of difficulty in word learning. Additional studies linked alpha desynchronization to cognitive effort and cognitive load[46,47]. So, one could assume to observe higher alpha desynchronization in the more difficult to learn condition of low-PP compared to high-PP. On the other hand numerous studies investigating oscillatory correlates of learning and memory showed that alpha desynchronization is associated with memory across different tasks, modalities and experimental phases of encoding and retrieval[35,36,44,45,48,49]. Strikingly, Griffith and colleagues[49] revealed by simultaneous EEG-fMRI recordings a negative correlation between the occurrence of patterns of stimulus-specific information detected by fMRI and cortical alpha/beta suppression. Here, the authors suggested that a decrease of alpha/beta oscillations might represent the neuronal mechanism of unmasking the task-critical signal by simultaneous suppression of task-irrelevant neuronal activities to promote information processing. Following this interpretation, we assume that over the course of learning elevated memory processing of the easier to learn stimuli is associated with enhanced information processing and thus accompanied by higher cortical alpha desynchronization in comparison of the more difficult to learn stimuli.

Our study has limitations due to a small sample size and between-subject comparisons. The criteria of data analyses were not pre-registered and the *p*-values of our behavior analyses were not corrected for multiple comparisons. Further, potential influences of phonotactic word properties and the prior encoding depth on TMR cannot be interpreted independently because the conditions of phonotactic word properties impacted already the encoding depth before sleep. To disentangle these potential factors, future studies could include a higher number of encoding trials for the low-PP words to equalize the encoding depth between conditions prior to TMR, as successfully shown in a recent study[20]. Additionally, a within-subject design with a larger sample size would provide a more robust control of interindividual differences in sleep and cognition. To control for potential confounders despite the influence of difficulty in word

learning on TMR, we compared parameters of sleep, the pre-sleep memory performance and the vigilance shortly before the post-sleep memory test, revealing no significant group differences (see Table S1 and S2). Nevertheless, we cannot rule out that other individual trait factors differed between the groups, such as the individual susceptibility to TMR. To rule out these alternative explanations based on individual factors, we suggest for future research to replicate our study by conducting a within-subject design with cueing of subsets of previously learned low- and high-PP words providing all conditions within the same individuals as shown in other TMR studies[25,26]. Further, we used artificial words based on American English in combination with German speaking participants, whereas language differences of pronunciation and phoneme structures might affect word perception and memory processing[50]. On the other hand, both languages are considered to have the same language family[51] and the phonological distance between English and German is quite short compared for example to Korean[52]. Thus, major common phonological characteristics across both languages are still preserved. In addition, our behavior analyses revealed robust word discrimination learning and distinct memory performance according to different levels of phonotactic probabilities providing evidence of successful experimental manipulation. Based on our paradigm, we investigated categorization learning of artificial words according to their reward associations (rewarded vs. unrewarded) and did not studied aspects of generalization learning of artificial grammar rules[53–55]. This difference might limit the comparability between these previous language-related studies and our findings. However, the usage of artificial words with distinct phonotactical properties provided a successful way to manipulate learning difficulty and to investigate word properties on TMR, whereas our reward categorization learning paradigm had the advantage to increase the relevance of the word learnings due to incentives.

To conclude, our present study demonstrates that difficulty in word learning as a manipulation of phonotactic properties impacts the effectiveness of TMR. Here, TMR of the easy to learn words facilitates categorization learning while TMR of the difficult words had no effect. Auditory cueing of the easy words induced increased spindle activity during the SW up-states - as an assumed underlying activity of memory reactivation - compared to the difficult words. Thus, our findings suggest that future research and clinical applications to restore language capabilities should consider learning difficulty based on phonotactical word properties as a potential factor of successful TMR, whereas alpha desynchronization during initial learning - as a neural marker of encoding depth - may serve as a predictor for sleep-associated memory reactivation and consolidation.

## References


1. Frisch, S. A., Large, N. R. & Pisoni, D. B. Perception of Wordlikeness: Effects of Segment Probability and Length on the Processing of Nonwords. *J. Mem. Lang.* **42**, 481–496 (2000).

2. Gupta, P. & Tisdale, J. Word learning, phonological short-term memory, phonotactic probability and long-term memory: Towards an integrated framework. *Philos. Trans. R. Soc. B Biol. Sci.* **364**, 3755–3771 (2009).

3. Gilboa, A. & Marlatte, H. Neurobiology of Schemas and Schema-Mediated Memory. *Trends Cogn. Sci.* **21**, 618–631 (2017).

4. Davis, M. H. & Gaskell, M. G. A complementary systems account of word learning: Neural and behavioural evidence. *Philos. Trans. R. Soc. B Biol. Sci.* **364**, 3773–3800 (2009).

5. McClelland, J. L., McNaughton, B. L. & O'Reilly, R. C. Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychol. Rev.* **102**, 419–457 (1995).

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

7. Mölle, M., Bergmann, T. O., Marshall, L. & Born, J. Fast and slow spindles during the sleep slow oscillation: disparate coalescence and engagement in memory processing. *Sleep* **34**, 1411–21 (2011).

8. Cox, R., van Driel, J., de Boer, M. & Talamini, L. M. Slow oscillations during sleep coordinate interregional communication in cortical networks. *J. Neurosci.* **34**, 16890–901 (2014).

9. Staresina, B. P. *et al.* Hierarchical nesting of slow oscillations, spindles and ripples in the human hippocampus during sleep. *Nat. Neurosci.* **18**, 1679–1686 (2015).

10. Siapas, A. G. & Wilson, M. A. Coordinated Interactions between Hippocampal Ripples and Cortical Spindles during Slow-Wave Sleep. *Neuron* **21**, 1123–1128 (1998).

11. Sirota, A. & Buzsáki, G. Interaction between neocortical and hippocampal networks via slow oscillations. *Thalamus Relat. Syst.* **3**, 245–259 (2005).

12. Diekelmann, S. & Born, J. The memory function of sleep. *Nat. Rev. Neurosci.* **11**, 114–126 (2010).

13. Buzsáki, G. Two-stage model of memory trace formation: A role for 'noisy' brain states. *Neuroscience* **31**, 551–570 (1989).

14. Girardeau, G. & Zugaro, M. Hippocampal ripples and memory consolidation. *Curr. Opin. Neurobiol.* **21**, 452–9 (2011).

15. Singer, A. C., Carr, M. F., Karlsson, M. P. & Frank, L. M. Hippocampal SWR activity predicts correct decisions during the initial learning of an alternation task. *Neuron* **77**, 1163–73 (2013).

16. Havas, V. *et al.* Semantic and phonological schema influence spoken word learning and overnight consolidation. *Q. J. Exp. Psychol.* **71**, 1469–1481 (2018).

17. Payne, J. D. *et al.* Memory for semantically related and unrelated declarative information: The benefit of sleep, the cost of wake. *PLoS One* **7**, (2012).

18. Zion, D. Ben, Nevat, M., Prior, A. & Bitan, T. Prior Knowledge Predicts Early Consolidation in Second Language Learning. *Front. Psychol.* **10**, (2019).

19. Durrant, S. J., Cairney, S. A., McDermott, C. & Lewis, P. A. Schema-conformant memories are preferentially consolidated during REM sleep. *Neurobiol. Learn. Mem.* **122**, 41–50 (2015).

20. Cordi, M. J., Schreiner, T. & Rasch, B. Is prior knowledge essential? Additional training opportunities restore sleep-associated memory benefits under conditions of low prior knowledge. *J. Sleep Res.* **32**, 1–9 (2023).

21. Rasch, B., Büchel, C., Gais, S. & Born, J. Odor cues during slow-wave sleep prompt declarative memory consolidation. *Science.* **315**, 1426–1429 (2007).

22. Diekelmann, S., Büchel, C., Born, J. & Rasch, B. Labile or stable: opposing consequences for memory when reactivated during waking and sleep. *Nat. Neurosci.* **14**, 381–386 (2011).

23. Diekelmann, S., Biggel, S., Rasch, B. & Born, J. Offline consolidation of memory varies with time in slow wave sleep and can be accelerated by cuing memory reactivations. *Neurobiol. Learn. Mem.* **98**, 103–111 (2012).

24. Rudoy, J. D., Voss, J. L., Westerberg, C. E. & Paller, K. A. Strengthening individual memories by reactivating them during sleep. *Science.* **326**, 1079 (2009).

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

26. Cairney, S. A., Guttesen, A. á. V., El Marj, N. & Staresina, B. P. Memory Consolidation Is Linked to Spindle-Mediated Information Processing during Sleep. *Curr. Biol.* **28**, 948-954.e4 (2018).

27. Antony, J. W. *et al.* Sleep Spindle Refractoriness Segregates Periods of Memory Reactivation. *Curr. Biol.* **28**, 1736-1743.e4 (2018).

28. Groch, S., Schreiner, T., Rasch, B., Huber, R. & Wilhelm, I. Prior knowledge is essential for the beneficial effect of targeted memory reactivation during sleep. *Sci. Rep.* **7**, (2017).

29. Creery, J. D., Oudiette, D., Antony, J. W. & Paller, K. A. Targeted Memory Reactivation during Sleep Depends on Prior Learning. *Sleep* **38**, 755–63 (2015).

30. Asfestani, M. A. *et al.* Consolidation of Reward Memory during Sleep Does Not Require Dopaminergic Activation. *J. Cogn. Neurosci.* **32**, 1688–1703 (2020).

31. Fischer, S. & Born, J. Anticipated reward enhances offline learning during sleep. *J. Exp. Psychol. Learn. Mem. Cogn.* **35**, 1586–1593 (2009).

32. Lansink, C. S., Goltstein, P. M., Lankelma, J. V., McNaughton, B. L. & Pennartz, C. M. A. Hippocampus leads ventral striatum in replay of place-reward information. *PLoS Biol.* **7**, (2009).

33. Sterpenich, V. *et al.* Reward biases spontaneous neural reactivation during sleep. *Nat. Commun. 2021 121* **12**, 1–11 (2021).

34. Green, D. M. & Swets, J. A. *Signal detection theory and psychophysics. Signal detection theory and psychophysics.* (John Wiley, 1966).

35. Hanslmayr, S., Spitzer, B. & Bäuml, K.-H. Brain oscillations dissociate between semantic and nonsemantic encoding of episodic memories. *Cereb. Cortex* **19**, 1631–40 (2009).

36. Griffiths, B. J., Martín-Buro, M. C., Staresina, B. P., Hanslmayr, S. & Staudigl, T. Alpha/beta power decreases during episodic memory formation predict the magnitude of alpha/beta power decreases during subsequent retrieval. *Neuropsychologia* **153**, (2021).

37. James, E., Gaskell, M. G., Weighall, A. & Henderson, L. Consolidation of vocabulary during sleep: The rich get richer? *Neurosci. Biobehav. Rev.* **77**, 1–13 (2017).

38. Ai, S. *et al.* Promoting subjective preferences in simple economic choices during nap. *Elife* **7**, (2018).

39. Klinzing, J. G., Niethard, N. & Born, J. Mechanisms of systems memory consolidation during sleep. *Nat. Neurosci.* **22**, 1598–1610 (2019).

40. Tamminen, J., Payne, J. D., Stickgold, R., Wamsley, E. J. & Gaskell, M. G. Sleep spindle activity is associated with the integration of new memories and existing knowledge. *J. Neurosci.* **30**, 14356–60 (2010).

41. Tamminen, J., Lambon Ralph, M. A. & Lewis, P. A. The role of sleep spindles and slow-wave activity in integrating new information in semantic memory. *J. Neurosci.* **33**, 15376–15381 (2013).

42. Schreiner, T., Petzka, M., Staudigl, T. & Staresina, B. P. Endogenous memory reactivation during sleep in humans is clocked by slow oscillation-spindle complexes. *Nat. Commun.* **12**, (2021).

43. Klimesch, W., Doppelmayr, M. & Hanslmayr, S. Chapter 10 Upper alpha ERD and absolute power: their meaning for memory performance. *Progress in Brain Research* vol. 159 151–165 (2006).

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

45. Griffiths, B. J. *et al.* Directional coupling of slow and fast hippocampal gamma with neocortical alpha/beta oscillations in human episodic memory. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 21834–21842 (2019).

46. Proskovec, A. L., Heinrichs-Graham, E. & Wilson, T. W. Load Modulates the Alpha and Beta Oscillatory Dynamics Serving Verbal Working Memory. *Neuroimage* **184**, 256 (2019).

47. Zhu, Y., Wang, Q. & Zhang, L. Study of EEG characteristics while solving scientific problems with different mental effort. *Sci. Rep.* **11**, (2021).

48. Griffiths, B., Mazaheri, A., Debener, S. & Hanslmayr, S. Brain oscillations track the formation of episodic memories in the real world. *Neuroimage* **143**, 256–266 (2016).

49. Griffiths, B. J. *et al.* Alpha/beta power decreases track the fidelity of stimulus specific information. *Elife* **8**, (2019).

50. Bohn, O. S. & Best, C. T. Native-language phonetic and phonological influences on perception of American English approximants by Danish and German listeners. *J. Phon.* **40**, 109–128 (2012).

51. Eberhard, D. M., Simons, G. F. & Fennig, C. D. Ethnologue: Languages of the world. SIL International. *Online version: http://www.ethnologue.com* (2019).

52. Luef, E. M. & Resnik, P. Phonotactic Probabilities and Sub-syllabic Segmentation in Language Learning. *Theory Pract. Second Lang. Acquis.* **9**, 1–31 (2023).

53. Reber, A. S. Implicit learning of artificial grammars. *J. Verbal Learning Verbal Behav.* **6**, 855–863 (1967).

54. Batterink, L. J. & Paller, K. A. Sleep-based memory processing facilitates grammatical generalization: Evidence from targeted memory reactivation. *Brain Lang.* **167**, 83–93 (2017).

55. Batterink, L. J., Oudiette, D., Reber, P. J. & Paller, K. A. Sleep facilitates learning a new linguistic rule. *Neuropsychologia* **65**, 169–79 (2014).

56. Aljasser, F. & Vitevitch, M. S. A Web-based interface to calculate phonotactic probability for words and nonwords in Modern Standard Arabic. *Behav. Res. Methods* **50**, 313–322 (2018).

57. Dinges, D. F. & Powell, J. W. Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behav. Res. Methods, Instruments, Comput.* **17**, 652–655 (1985).

58. Macmillan, N. A. & Kaplan, H. L. Detection Theory Analysis of Group Data. Estimating Sensitivity From Average Hit and False-Alarm Rates. *Psychol. Bull.* **98**, 185–199 (1985).

59. Iber, C., Ancoli-Israel, S., Chesson, A. L. & Quan, S. F. *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*. (American Academy of Sleep Medicine, 2007).

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

## Methods

Subjects

The study included 39 subjects (29 females) with an age range of 19 to 28 years (M = 22.28 years, SD = 2.04). Participants were recruited from the University of Fribourg community by E-Mail or through advertisements at the campus of the University. Before participation, subjects had to give written informed consent as approved by the Ethical Commission of the Department of Psychology of the University of Fribourg. All participants were German speakers, and no subject had a history of neurological or psychiatric illness. The participants were instructed to keep a normal sleep schedule, to get up in the morning before 8 a.m. and not to consume alcohol and caffeine on experimental days. For participation, subjects either received credit for an undergraduate class and/or monetary compensation.
Data were excluded from subjects who did not reach the minimal learning performance of d' > 1.05 during the pre-sleep memory test in at least one of the two PP conditions, whereas this threshold value corresponds to accuracy rates of 70% ($n$ = 5). In addition, we excluded one subject who showed a negative d' in one PP condition of the pre-sleep memory test ($n$ = 1).

Pre-Learning

Participants arrived at the sleep laboratory at 19:30h. Electrodes for standard polysomnography (32 EEG electrodes, EMG and ECG electrodes) were mounted with two EEG mastoid electrodes. One electrode under the right eye was attached to record eye movements.

Encoding task

The encoding task started around 21:00h. Subjects learned to discriminate rewarded and unrewarded artificial words by right- and left-hand button presses. As a two alternative forced-choice task, we assigned left- and right-hand button presses to the rewarded and the unrewarded word category, counterbalanced across subjects. We instructed the participants to respond to each word by left- or right-hand button presses, whereas one button means the word is rewarded (gain of money points) and the other button means the word is unrewarded (avoid the loss of money points). During each trial, an initial fixation cross was displayed between 1.7 and 2.3s. Subsequently, concurrent with the onset of a blank screen the sound of an artificial word was played for 0.7 to 1s. After a waiting period (blank screen 2s), a question mark appeared to signal the onset of the response time window (maximal duration 4s). Following the key press response, a feedback screen with the money points of the trial and the current task score appeared for 2s. The learning task contained 480 trials in randomized order according to 160 artificial words with three presentations each.

Pre-Sleep Memory task

After finishing the encoding task, following a 10min break, subjects performed the pre-sleep memory task. The memory task had the same trial structure as the encoding task without the last feedback screen. Participants received again money points for correct responses. However, the received amount was only shown at the end of the task. The memory task contains 160 trials according to the 160 artificial words.

Auditory target memory reactivation during sleep

Following an additional impedance check and re-adjustment of the EEG electrodes, subjects went to bed in a noise and electric shielded cabin of the sleep laboratory. All night sleep periods

started with light off between 23:00 and 24:00h. Based on online monitoring of N2 and N3 sleep, artificial words were presented aurally via loudspeakers (sound pressure level 55dB) with a randomized inter-stimulus interval of 8±2s. One group of subjects were exposed to artificial words with low phonotactic probability during sleep, whereas we presented to the other group the high phonotactic probability words (see Table S1 for number of reactivations). We interrupted word presentations when we observed online an arousal or patterns of REM sleep.

Post-Sleep Memory task

After sleep and re-adjustment of the EEG, subjects performed the post-sleep memory task (see above the description of the pre-sleep memory task).

Experimental tasks

All experimental tasks, including sleep reactivation, were conducted by using the E-Prime software (Psychology Software Tools, Sharpsburg, USA). By presentation of practice trials at the beginning, the subject's understanding of the task was approved. Across all tasks, stimuli were presented in randomized order.

Stimuli

To create artificial words, we subdivided the alphabet in two groups of consonants (C1: b, c, d, f, g, h, j, k, l, m; C2: n, p, q, r, s, t, v, w, x, z) and vowels (V1: a, e, I; V2: o, u, y). Four-letter-words were created by selecting letters from the vowel and consonant groups according to four different sequences (G1:C1, V1, V2, C2; G2: C1, V1, C2, V2; G3: V1, C1, C2, V2; G4: V1, C1, V2, C2). Artificial words were converted automatically from text to speech files (wav format) by using MATLAB functions (tts.m, audiowrite.m) with the setting of a female computer voice with American English pronunciation. From this pool of 900 stimuli for each rule, we selected 40 artificial words per rule category (4 × 40 words) according to the inclusion criteria: fluent and understandable pronunciation; no meaning; no names; aurally discriminable to other selected words (see Table S3 and S4 for lists of the used words).
Phonotactic probability values were computed for each word by using an online computation platform[56] (https://calculator.ku.edu/phonotactic/about) according to an American English lexicon. Between rule category comparison analyses revealed significant differences in phonotactic probabilities between two rules in contrast to the other two rules (see Fig. 1b).

Psychomotor vigilance task

With respect to vigilance after sleep before the post-sleep memory task, a reaction time measurement was conducted by the Psychomotor Vigilance Task (PVT)[57]. Each trial began with a fixation cross of a randomized duration between 2 and 10 seconds. Participants were instructed to press the spacebar button with the forefinger of the non-dominant hand as quickly as possible after numbers started to count in milliseconds on the screen. The reaction time was displayed for 1s after the key press. The PVT had a duration of 10 min (see for results Table S2).

Behavior analyses

As a measurement of discrimination learning and memory performance, we calculated d'-values (d'=z(hits)-z(false alarms)) according to the signal detection theory[34]. Here, we considered correct trials with gaining of money points as hits and incorrect trials with a loss of money points as false alarms (see Fig. 1d). In addition, we assessed measurements of the response bias

(c=-0.5*(z(false alarms)+z(hits))). According to the convention[58], rates of 0 and 1 were replaced with 0.5/*n* and with (*n*-0.5)/*n* respectively, whereas *n* corresponds to the number of trials (*n* = 40). Values of d' and response bias were calculated separately for the low- and high-PP condition.

EEG recordings

We made EEG recordings by using customized 32-Ag/AgCl electrodes at 10-10 locations caps (EASYCAP, Woerthsee-Etterschlag, Germany) and 32 channel amplifiers (Brain Products, Gilching, Germany). Impedances were kept below 10 kΩ. The EEG was recorded with a sampling rate of 500Hz using Brain Vision Recorder software (Brain Products, Gilching, Germany). Signals were referenced to electrodes at the mastoids. The ocular activity was measured via one EOG channel mounted ~2 cm below the right eye. Muscle tone was monitored by EMG recordings made under the chin. The following EEG electrodes were used for subsequent wake and sleep analyses: Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T7, T8, P7, P8, Fz, Cz and Pz.

Sleep scoring

The sleep stages of NREM 1 to 3 (N1 to N3), wake, and REM sleep were scored offline and manually according to the criteria of the American Academy of Sleep Medicine (AASM) by visual inspection of the signals of the frontal, central, and occipital electrodes over 30s epochs[59]. Based on offline scoring, we confirmed TMR exposure during N2 and N3 and no significant differences (*p*-values > 0.05) of sleep parameters between the cueing groups (see Table S2).

Preprocessing Wake EEG

Wake EEG were preprocessed using the Fieldtrip toolbox (http://fieldtriptoolbox.org; Donders Institute for Brain, Cognition and Behaviour, Radboud University, Netherlands)[60]. First, raw data signals were re-referenced to averaged mastoid electrodes and filtered (high- and low-pass, 0.5 and 45Hz). For each trial, data were segmented into epochs (-2 to 4s) time-locked to stimulus onset. After demean and detrend processing steps, noisy trials were identified by visual inspection and discarded from further analyses. We then conducted an independent component analysis (ICA) to identify and reject ICA components impacted by eye blinks and eye movements.

Preprocessing Sleep EEG

EEG signals during sleep were re-referenced to averaged mastoid electrodes and filtered (high- and low-pass, 0.5 and 35Hz). To analyze EEG activities of TMR, data were segmented into epochs (-2 to 8s) around the onset of stimuli presentations. After demean and detrend processing, trials with signals distorted by noise and movement artefacts were identified by visual inspection and removed from further analyses.

Time-Frequency Analysis Wake data

We conducted the time-frequency analyses by using Morlet wavelet analyses implemented in the Fieldtrip toolbox. The number of wavelet cycles was adjusted to 7 cycles. We extracted oscillatory power in the frequency range of 1 to 45Hz with frequency steps of 0.2Hz and time steps of 10ms. All the power values were baseline corrected and transformed to absolute

changed values by subtracting the corresponding averaged values of the baseline interval of -1 to -0.1s before stimulus onset.

Slow wave-spindle power analyses

To reduce influence of potential confounders on TMR like low prior learning performance and number of reactivations, we included only subjects with a pre-sleep memory performance d' > 0.75 in both PP conditions and who did receive a minimal number of 160 reactivations (each word was represented at least two times during sleep). Thus, we excluded 5 and 6 subjects respectively. The final sample for sleep analyses consisted of $n$ = 22 with 11 subjects in each cueing condition.

To detect slow waves, we first localized all negative and positive peaks of the 0.5 to 3Hz band-pass filtered signal during the auditory cueing trials in the time windows of 0 to 6s according to stimulus onsets. In addition, time distance between the prior and posterior positive peaks should be in the range of 0.3 to 2s. For the following analyses, the 30% slow waves with the highest amplitude (from negative peak to posterior positive peak) were included for each subject.

We then extracted the corresponding spectral power (1-30Hz) in a time window around this negative slow wave peak (±2s) using Morlet wavelet (7 cycles) analysis implemented in the Fieldtrip toolbox, with a time window and step size of 10ms and frequency steps of 0.2Hz. Extracted power data over frequency and time was averaged of the slow waves and baseline corrected by subtracting the averaged values of the baseline interval of -2 to -1.5s prior to the negative slow wave peak. Finally, the transformed absolute changed values were plotted for the frequency and time range of interest and different conditions (see Fig. 4c).

Slow wave density analyses

To analyze the number of slow waves around memory reactivations during sleep, we divided the time course of reactivation trials in equally sized bins with a duration of 0.5s from -1.5 to 3s according to stimulus onset. Based on negative peaks of detected slow waves, we counted the occurring of slow waves for each bin over trials. Subsequently, we computed density values for each bin by division of the number of trials. Finally, we calculated the relative increase in slow wave density in percent by division of the averaged baseline values (bins: -1.5 to 0s; see Fig. 4b). In an additional histogram, we show the grand average slow wave density distribution in number per trial without baseline correction and with a bin size of 0.5s from -0.5 to 6s time-locked to stimulus onset (see Fig. S7).

Sleep spindle analyses

We detected fast sleep spindles by band-pass filtering (12-16Hz) the signal of the Pz electrode during the auditory cueing trials in the time windows of -2 to 8s according to stimulus onsets. The amplitude threshold was calculated individually for each subject as 1.25 standard deviations (SDs) from the mean. The beginning and end times of the sleep spindles were then defined as the points at which the amplitude fell below 0.75 SDs before and after the detected sleep spindle. Only sleep spindles with a duration of 0.5-3 s were included in subsequent analyses.

To compare the sleep spindle densities between the different cueing conditions of high- and low-PP, we computed the grand average sleep spindle density distribution in number per trial

with a bin size of 0.5s from -0.5 to 6s time-locked to stimulus onset in each condition (see Fig. S8a and Table S9).

Based on the detected slow waves and sleep spindles, we defined coupling events when the positive amplitude peak of a detected sleep spindle was occurring during the slow wave up-state phase in a time window of 0.3 to 0.8s according to the trough of a slow wave.

We computed the averaged amplitude size of each detected sleep spindle by calculating the mean of the absolute amplitude values of all negative and positive peaks within a detected sleep spindle (see Fig. S8b).

Statistical Analysis

To test for differences between conditions of wake behavioral data, repeated-measure ANOVAs and paired sample *t*-tests were used. To analyze overnight changes of sleep behavioral data within TMR conditions, we conducted at first dependent sample *t*-tests against 0 of Δ-values (post-sleep test minus pre-sleep test) of d' and c-criterion (see Fig. 3). Two-way mixed design ANOVAs were computed to compare Δ-values between TMR conditions. After confirming at least a trend of significance ($p < 0.1$) for the interaction effect, we conducted post-hoc pairwise comparisons by independent and dependent sample *t*-tests. For all behavior statistical analyses, the *p*-value was set at $p < 0.05$ for two-tailed testing. A *p*-value < 0.1 and > 0.05 was reported as a trend of significance.

All statistical analyses were processed by using MATLAB 2018a (MathWorks, Natick, USA) and SPSS (IBM Corp., Version 25).

Considering the problem of multiple testing in statistics, we applied cluster based permutation tests implemented in the Fieldtrip toolbox[60] to test for potential differences of EEG time-frequency data. Here, dependent sample *t*-tests and an independent sample test (for between subject group comparison of sleep data) were conducted. We specified the number of permutations by 1000 and the minimal number of channels by two. The cluster alpha level was set at $p = 0.05$ for two-tailed testing. Plots with significant clusters of electrodes are shown in Fig. 2d and 4d. After confirming the existence of a significant cluster, we conducted an additional post-hoc repeated-measure ANOVA with averaged values of the identified time and frequency range of interest and merged over the Pz and Cz electrodes (see Fig. 2e). To control for SW amplitude size as a potential confounder in spindle activity differences during SW up-states, we extracted and averaged power values of the significant cluster (0.3 to 0.8s; 11-14Hz; Pz, P3, P4, O2, P7) per subject and conducted a post-hoc independent sample *t*-test based on residuals corrected by individual SW amplitude size. By using the same extracted power values (0.3 to 0.8s; 11-14Hz; Pz, P3, P4, O2, P7) per subject, we performed whole sample ($n = 22$) Pearson correlation analyses between these power values and the overnight changes of behavior measurements of the cued condition (Δ d' and Δ c-criterion).

**Supplementary data**

**Supplementary table S1** Gender, age and performance rates

| | High-PP cued (*n* = 11) | Low-PP cued (*n* = 11) | *t* | *P* |
|---|---|---|---|---|
| Female (#) | 10 | 7 | - | - |
| Age (years) | 21.91 ± 0.33 | 22.00 ± 0.56 | -0.14 | 0.89 |
| Pre-sleep memory test | | | | |
| High-PP (d') | 2.11 ± 0.16 | 2.03 ± 0.26 | 0.27 | 0.79 |
| Low-PP (d') | 1.58 ± 0.16 | 1.72 ± 0.23 | -0.53 | 0.60 |
| High-PP (c-criterion) | 0.30 ± 0.09 | 0.57 ± 0.11 | -1.86 | 0.08 |
| Low-PP (c-criterion) | 0.34 ± 0.09 | 0.60 ± 0.14 | -1.60 | 0.13 |
| Psychomotor vigilance task before post-sleep memory test | | | | |
| Reaction time (ms) | 319.97 ± 8.99 | 333.63 ± 10.19 | -1.01 | 0.33 |

Data are means ± SEM. PP, phonotactic probability. *P*-values of statistical comparisons between groups by using unpaired *t*-tests. Note, no significant group differences, but a trend of significance for pre-sleep memory test of c-criterion values of the high-PP condition.

**Supplementary table S2** Sleep and reactivation parameter

| | High-PP cued ($n$ = 11) | Low-PP cued ($n$ = 11) | $t$ | $P$ |
|---|---|---|---|---|
| Duration (min) | | | | |
| N1 | 91.55 ± 20.02 | 65.64 ± 18.77 | 0.94 | 0.36 |
| N2 | 181.64 ± 11.61 | 200.55 ± 14.5 | -1.02 | 0.32 |
| N3 | 119.96 ± 15.11 | 99.46 ± 12.11 | 1.06 | 0.30 |
| REM | 57.91 ± 13.30 | 88.86 ± 9.87 | -1.87 | 0.08 |
| WASO | 16.96 ± 10.18 | 5.55 ± 1.19 | 1.11 | 0.29 |
| Duration (%) | | | | |
| N1 | 19.44 ± 4.18 | 13.98 ± 3.89 | 0.96 | 0.35 |
| N2 | 38.85 ± 2.56 | 43.77 ± 3.30 | -1.18 | 0.25 |
| N3 | 25.54 ± 3.16 | 21.55 ± 2.52 | 0.99 | 0.34 |
| REM | 12.40 ± 2.82 | 19.46 ± 2.20 | -1.98 | 0.06 |
| WASO | 3.71 ± 2.27 | 1.23 ± 0.27 | 0.28 | 0.29 |
| Number of reactivations | | | | |
| N2 | 61.00 ± 22.22 | 93.18 ± 24.62 | -0.97 | 0.34 |
| N3 | 266.18 ± 53.68 | 231.73 ± 34.43 | 0.54 | 0.60 |
| Slow wave detection | | | | |
| SW density (#/3s) | 1.87 ± 0.09 | 1.99 ± 0.13 | -0.76 | 0.46 |
| SW amplitude (µV) | 134.33 ± 14.28 | 107.78 ± 9.96 | 1.53 | 0.14 |

Data are means ± SEM. N1, N2: Non-REM sleep stages N1, N2 and N3; REM, rapid-eye movement sleep; WASO, wake after sleep onset; SW, slow wave; PP, phonotactic probability. *P*-values of statistical comparisons between groups by using unpaired *t*-tests. Note, no significant group differences, but a trend of significance for REM sleep parameters.

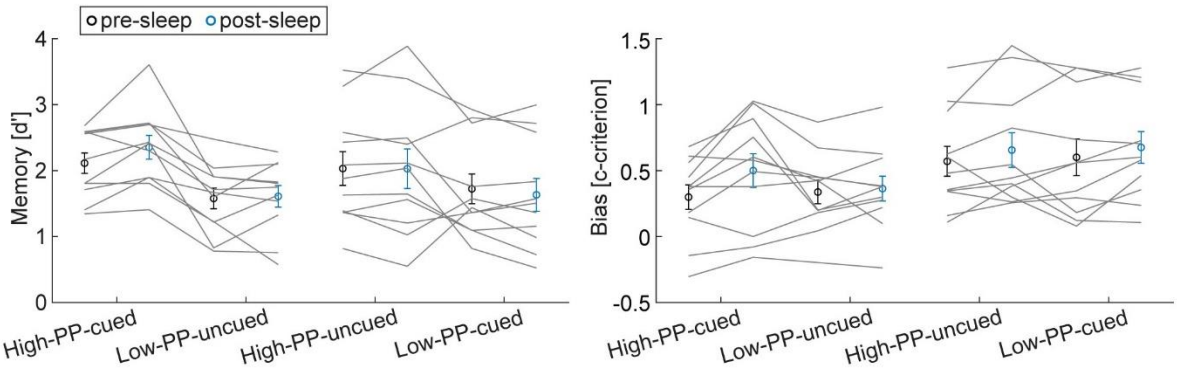
**Supplementary table S3** List of the high-PP words and their phoneme and biphone probabilities

| | *G1 (High-PP)* | *Phonotactic Probabilities* | | *G2 (High-PP)* | *Phonotactic Probabilities* | |
|---|---|---|---|---|---|---|
| ***Number*** | ***Word*** | ***Phoneme*** | ***Biphone*** | ***Word*** | ***Phoneme*** | ***Biphone*** |
| 1 | baos | 0.180 | 0.005 | baqy | 0.126 | 0.004 |
| 2 | beon | 0.146 | 0.003 | basy | 0.205 | 0.006 |
| 3 | biox | 0.181 | 0.003 | bavo | 0.156 | 0.005 |
| 4 | biuq | 0.098 | 0.002 | beto | 0.167 | 0.006 |
| 5 | cayr | 0.114 | 0.000 | bevo | 0.125 | 0.004 |
| 6 | ceot | 0.145 | 0.002 | bisu | 0.173 | 0.004 |
| 7 | ceov | 0.070 | 0.001 | bitu | 0.160 | 0.005 |
| 8 | ceox | 0.135 | 0.000 | bixy | 0.121 | 0.003 |
| 9 | cion | 0.105 | 0.002 | catu | 0.145 | 0.004 |
| 10 | cioz | 0.070 | 0.001 | cipu | 0.088 | 0.002 |
| 11 | daon | 0.178 | 0.004 | devo | 0.126 | 0.004 |
| 12 | daot | 0.220 | 0.004 | disy | 0.177 | 0.005 |
| 13 | deop | 0.136 | 0.002 | fasu | 0.197 | 0.005 |
| 14 | deus | 0.146 | 0.002 | fato | 0.194 | 0.006 |
| 15 | diov | 0.117 | 0.004 | faxo | 0.152 | 0.002 |
| 16 | faop | 0.162 | 0.002 | gazu | 0.118 | 0.003 |
| 17 | faoq | 0.126 | 0.002 | gexo | 0.101 | 0.002 |
| 18 | faor | 0.169 | 0.007 | giny | 0.168 | 0.006 |
| 19 | faot | 0.215 | 0.004 | haqy | 0.114 | 0.004 |
| 20 | faox | 0.206 | 0.002 | haso | 0.200 | 0.007 |
| 21 | faoz | 0.138 | 0.002 | heso | 0.168 | 0.005 |
| 22 | gaor | 0.148 | 0.007 | hesu | 0.159 | 0.004 |
| 23 | geon | 0.121 | 0.003 | hevo | 0.113 | 0.003 |
| 24 | giur | 0.115 | 0.000 | jaqo | 0.081 | 0.000 |
| 25 | gius | 0.123 | 0.001 | jiry | 0.124 | 0.000 |
| 26 | giut | 0.162 | 0.002 | kapy | 0.204 | 0.021 |
| 27 | haor | 0.161 | 0.008 | kexu | 0.158 | 0.002 |
| 28 | heoq | 0.087 | 0.001 | kipo | 0.183 | 0.004 |
| 29 | heox | 0.167 | 0.001 | kipu | 0.173 | 0.002 |
| 30 | hion | 0.137 | 0.003 | kiqo | 0.146 | 0.001 |
| 31 | jaoq | 0.079 | 0.000 | kizo | 0.166 | 0.002 |
| 32 | jios | 0.101 | 0.002 | leqo | 0.084 | 0.002 |
| 33 | kaon | 0.219 | 0.018 | lexu | 0.099 | 0.002 |
| 34 | kaot | 0.261 | 0.018 | liwo | 0.092 | 0.003 |
| 35 | kaov | 0.187 | 0.017 | manu | 0.225 | 0.019 |
| 36 | kiuz | 0.151 | 0.001 | mawo | 0.144 | 0.006 |
| 37 | leon | 0.129 | 0.003 | meqy | 0.101 | 0.003 |
| 38 | leoq | 0.082 | 0.002 | mino | 0.206 | 0.006 |
| 39 | liux | 0.160 | 0.003 | mipy | 0.140 | 0.005 |
| 40 | miux | 0.184 | 0.002 | mivy | 0.127 | 0.004 |

**Supplementary table S4** List of the low-PP words and their phoneme and biphone probabilities

| | *G3 (Low-PP)* | *Phonotactic Probabilities* | | *G4 (Low-PP)* | *Phonotactic Probabilities* | |
|---|---|---|---|---|---|---|
| ***Number*** | ***Word*** | ***Phoneme*** | ***Biphone*** | ***Word*** | ***Phoneme*** | ***Biphone*** |
| 1 | abpu | 0.069 | 0.002 | afut | 0.121 | 0.002 |
| 2 | abwo | 0.047 | 0.002 | afux | 0.112 | 0.000 |
| 3 | adro | 0.120 | 0.003 | agyv | 0.035 | 0.000 |
| 4 | adwo | 0.047 | 0.000 | ahon | 0.078 | 0.001 |
| 5 | adzo | 0.062 | 0.001 | ahun | 0.074 | 0.001 |
| 6 | afwo | 0.043 | 0.000 | ahuz | 0.039 | 0.001 |
| 7 | afzo | 0.058 | 0.000 | ajot | 0.120 | 0.002 |
| 8 | agno | 0.135 | 0.002 | alor | 0.118 | 0.008 |
| 9 | agvy | 0.055 | 0.000 | ebon | 0.078 | 0.002 |
| 10 | agzo | 0.059 | 0.002 | ebox | 0.111 | 0.001 |
| 11 | ahvy | 0.050 | 0.000 | ecyn | 0.070 | 0.000 |
| 12 | akvu | 0.074 | 0.001 | efon | 0.074 | 0.002 |
| 13 | alpo | 0.115 | 0.002 | ehos | 0.073 | 0.001 |
| 14 | alzu | 0.088 | 0.000 | ehyq | 0.007 | 0.000 |
| 15 | ampy | 0.083 | 0.009 | ehyt | 0.096 | 0.000 |
| 16 | ebry | 0.105 | 0.001 | ehyv | 0.022 | 0.000 |
| 17 | ebvy | 0.050 | 0.001 | ejyz | 0.019 | 0.000 |
| 18 | ecpy | 0.072 | 0.002 | ekoz | 0.062 | 0.002 |
| 19 | ecto | 0.107 | 0.003 | ekyr | 0.077 | 0.001 |
| 20 | edty | 0.092 | 0.000 | ekyz | 0.046 | 0.001 |
| 21 | edxu | 0.048 | 0.001 | elon | 0.114 | 0.005 |
| 22 | edxy | 0.051 | 0.001 | elyp | 0.088 | 0.000 |
| 23 | efny | 0.119 | 0.003 | elys | 0.102 | 0.000 |
| 24 | efqy | 0.023 | 0.000 | emos | 0.092 | 0.002 |
| 25 | efsy | 0.102 | 0.000 | emyn | 0.073 | 0.001 |
| 26 | efwy | 0.028 | 0.000 | ibon | 0.077 | 0.002 |
| 27 | egnu | 0.117 | 0.001 | icon | 0.085 | 0.001 |
| 28 | egro | 0.109 | 0.004 | icuq | 0.035 | 0.000 |
| 29 | egxy | 0.048 | 0.000 | idon | 0.077 | 0.002 |
| 30 | ekpo | 0.089 | 0.002 | idox | 0.110 | 0.000 |
| 31 | ibry | 0.104 | 0.000 | idoz | 0.043 | 0.001 |
| 32 | ibvu | 0.047 | 0.000 | idur | 0.069 | 0.000 |
| 33 | idno | 0.129 | 0.001 | ihos | 0.072 | 0.001 |
| 34 | idqo | 0.033 | 0.000 | ijut | 0.108 | 0.002 |
| 35 | igpy | 0.060 | 0.002 | ijyv | 0.021 | 0.000 |
| 36 | igru | 0.098 | 0.002 | iluq | 0.063 | 0.002 |
| 37 | igty | 0.089 | 0.000 | ilyt | 0.140 | 0.000 |
| 38 | igvo | 0.053 | 0.001 | ilyx | 0.131 | 0.000 |
| 39 | ikvu | 0.066 | 0.001 | imup | 0.074 | 0.002 |
| 40 | imno | 0.140 | 0.001 | imus | 0.087 | 0.001 |

## Supplementary figure S5



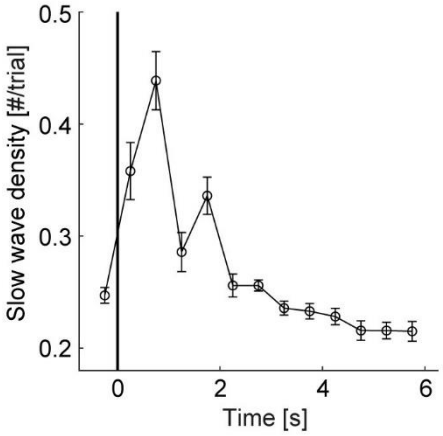

**Supplementary Fig. S5. Pre- and post-sleep memory test data.** Pre (black) and post (blue) sleep mean group memory test data of d' (left panel) and c-criterion (right panel) divided into the conditions of high- and low-PP and cued and uncued. Gray lines represent individual data points. Error bars reflect SEM.

**Supplementary table S6** Statistics of SW density time-bin analyses from -0.5 to 3s

| Time-bins (s) | SW density (%) | *t* | *P* |
|---|---|---|---|
| -0.5 to 0s | 0.21 ± 2.10 | 0.10 | 0.92 |
| 0 to 0.5s | 49.73 ± 14.32 | 3.47 | 0.002 |
| 0.5 to 1s | 83.44 ± 14.97 | 5.57 | <0.001 |
| 1 to 1.5s | 19.77 ± 10.26 | 1.93 | 0.07 |
| 1.5 to 2s | 40.33 ± 10.67 | 3.78 | 0.001 |
| 2 to 2.5s | 5.98 ± 6.37 | 0.94 | 0.36 |
| 2.5 to 3s | 4.81 ± 3.08 | 1.56 | 0.13 |

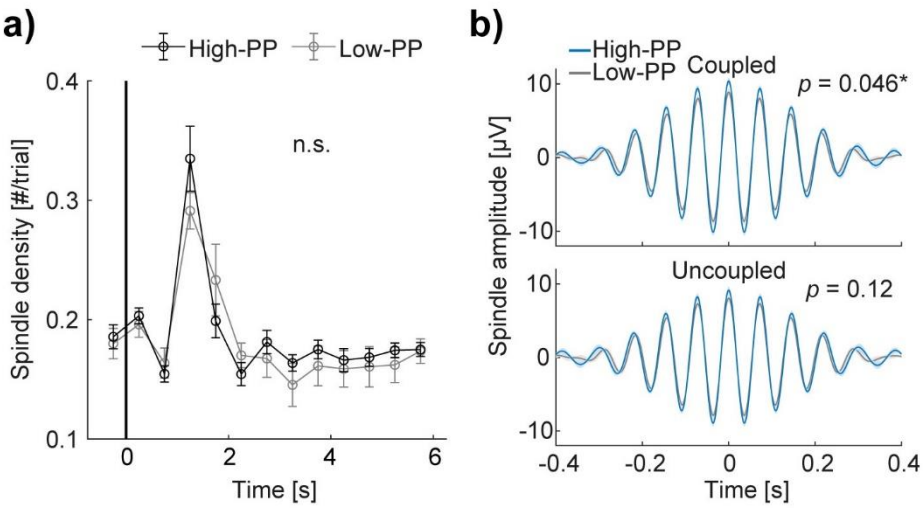
Data are means ± SEM. *P*-values (uncorrected for multiple comparisons) of statistical comparisons between groups by using paired *t*-tests against 0.

## Supplementary figure S7




**Supplementary Fig. S7. Slow wave density distribution**. Grand average curve of slow wave density time-locked to stimulus onset in number per trial with a bin size of 0.5s from -0.5 to 6s. Note, slow wave density peaks during the time window of 0 to 2s and lower density values from 3 to 6s after stimulus onset.

**Supplementary figure S8**




**Supplementary Fig. S8. Detected sleep spindle analyses.** a) Grand average curves of sleep spindle density time-locked to stimulus onset in number per trial with a bin size of 0.5s from -0.5 to 6s of the cueing condition of high (black) - and low (gray) - PP, respectively. Note, sleep spindle density peaks during the time window of 1 to 1.5s after stimulus onset with no significant (n.s., $p$>0.05) difference between conditions. b) Grand average signal filtered (12-16Hz) of the Pz electrode time-locked to the sleep spindle amplitude peak of the high (blue) - and low (gray) - PP cueing conditions; shaded areas represent SEM. Two different panels according to sleep spindles coupled (top) and uncoupled (bottom) with detected slow waves during the time window of 0 to 6s after stimulus onset. Note, significant higher amplitude sleep spindles coupled with slow waves in the cueing condition of high- compared to low-PP and no significant difference between the cueing conditions of the uncoupled sleep spindles. Error bars reflect SEM; *$p$<0.05.

After conducting a sleep spindle detection (frequency range of 12-16Hz, see methods for details), we compared the sleep spindle density between the TMR conditions of high- and low-PP showing no significant difference (see Fig. S8a and Table S9). Next, we subdivided the detected sleep spindles into coupled and uncoupled sleep spindles with the previously detected slow waves (SW; analyses of Fig. 4). Sleep spindles were defined as coupled when their amplitude peak occurred during the SW up-state phase (0.3 to 0.8s time-locked to the SW troughs). A two-way mixed design ANOVA on the amplitude size of the sleep spindles with the cueing group as a between-subject factor (high-PP-cued vs. low-PP-cued) and SW-coupling as a within-subject factor (coupled vs. uncoupled) showed a significant interaction effect (cueing group × SW-coupling: $F(1,20) = 4.51$, $p = 0.046$, $\eta^2 = 0.18$), a significant main effect of SW-coupling ($F(1,20) = 85.02$, $p < 0.001$, $\eta^2 = 0.81$), and a trend of significance of the main effect of the cueing group ($F(1,20) = 3.54$, $p = 0.08$). Post-hoc unpaired $t$-tests revealed a significant higher amplitude size of the coupled sleep spindles of the cueing group of high- compared to low-PP ($t(20) = 2.13$, $p = 0.046$, Cohen's $d = 0.91$; Fig. S8b) and no significant group difference of the uncoupled sleep spindles ($t(20) = 1.62$, $p = 0.12$). An additional comparison of the amount of coupled sleep spindles between the cueing groups revealed no significant difference (see Table S9).

Here, we found that detected sleep spindles coupled to the SW up-state phase occurred with higher amplitude after TMR presentations of the high-PP words in comparison to the low-PP words, whereas the sleep spindle density and the amount of sleep spindles coupled to the SW up-state phase did not differed between the cueing conditions.

**Supplementary table S9** Density parameters of detected sleep spindles

| | High-PP cued ($n$ = 11) | Low-PP cued ($n$ = 11) | $t$ | $P$ |
|---|---|---|---|---|
| Detected sleep spindles | | | | |
| General density (#/trial) | 2.25 ± 0.03 | 2.18 ± 0.11 | 0.60 | 0.56 |
| SW-coupled (%) | 38.10 ± 1.19 | 37.51 ± 2.32 | 0.22 | 0.83 |

Data are means ± SEM. PP, phonotactic probability; SW, slow wave. General density: sleep spindle density during the time window of 0 to 6s after stimulus onset in number per trial. SW-coupled: sleep spindles coupled to SW up-states divided by the total number of detected sleep spindles during the time window of interest (0 to 6s after stimulus onset) in percent. *P*-values of statistical comparisons between groups by using unpaired *t*-tests.

## Author contributions

Design of the study: A.-L.K. and B.R.; data acquisition: A-L.K.; statistical analyses and interpretation of the data: A.-L.K. and B.R.; writing the manuscript: A.-L.K. and B.R.

## Acknowledgements

We thank the students at the University of Fribourg who assisted in data collection and Dr. Jonas Beck for reading and providing comments to improve the manuscript. This work was supported by the Swiss National Science Foundation Grant 168602 to A.-L.K. and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement MemoSleep No. 677875) to B.R., and the University of Fribourg.

## Conflict of interest

The authors declare that they have no conflict of interest.

## Data availability statement

Data is available on request.