

Implicit memory for words heard during sleep

Thomas Andrillon^{1,2,*},† and Sid Kouider¹

¹Département d'Études Cognitives, École Normale Supérieure—PSL Research University, Brain and Consciousness Group (ENS, EHESS, CNRS), Paris, France; ²École Doctorale Cerveau Cognition Comportement, Université Pierre et Marie Curie, Paris, France

*Correspondence address: Département d'Études Cognitives, École Normale Supérieure—PSL Research University, Brain and Consciousness Group (ENS, EHESS, CNRS), Paris, France. Tel: +33 1 44 32 26 22; Fax: +33 1 44 32 26 30; E-mail: thomas.andrillon@gmail.com

†Andrillon, Thomas: <http://orcid.org/0000-0003-2794-8494>

Abstract

When we fall asleep, our awareness of the surrounding world fades. Yet, the sleeping brain is far from being dormant and recent research unraveled the preservation of complex sensory processing during sleep. In wakefulness, such processes usually lead to the formation of long-term memory traces, being it implicit or explicit. We examined here the consequences upon awakening of the processing of sensory information at a high level of representation during sleep. Participants were instructed to classify auditory stimuli as words or pseudo-words, through left and right hand responses, while transitioning toward sleep. An analysis of the electroencephalographic (EEG) signal revealed the preservation of lateralized motor activations in response to sounds, suggesting that stimuli were correctly categorized during sleep. Upon awakening, participants did not explicitly remember words processed during sleep and failed to distinguish them from new words (old/new recognition test). However, both behavioral and EEG data indicate the presence of an implicit memory trace for words presented during sleep. In addition, the underlying neural signature of such implicit memories markedly differed from the explicit memories formed during wakefulness, in line with dual-process accounts arguing for two independent systems for explicit and implicit memory. Thus, our results reveal that implicit learning mechanisms can be triggered during sleep and provide a novel approach to explore the neural implementation of memory without awareness.

Key words: sleep; memory; EEG; learning

Introduction

Sleepers are not disconnected from their environment during sleep. On the contrary, the sleeping brain can encode sensory information (Issa and Wang, 2008; Nir et al., 2013), recognize salient or familiar sounds such as a person's own name (Perrin et al., 1999), process sounds in their context and detect the violation of simple rules [oddball paradigm; (Czisch et al., 2009; Ruby et al., 2008; Strauss et al., 2015)]. Sleepers can also process sensory information at a high level of representation such as the semantic level (Bastuji et al., 2002; Brualla et al., 1998; Ibanez et al., 2006). We recently showed that the sleeping brain could even build upon sensory processes and use semantic information to prepare task-relevant responses (Kouider et al., 2014). But

can such processes trigger long-term memory? Indeed, when we are awake, experience constantly imprints on the brain. From pure noise (Andrillon et al., 2015) to more complex types of sensory inputs such as words (Pulvermüller et al., 2001), processing a given piece of information, even passively, leads to the formation of a new memory trace or the strengthening of an existing one (Kolb and Whishaw, 1998).

Exploring the sleeping brain's ability to learn has been a long-lasting scientific quest (Emmons and Simon, 1956) but positive results are scarce. Until recently, only some forms of learning independent from hippocampal structures had been evidenced [(Hennevin et al., 1995; Ikeda and Morotomi, 1996; Maho and Bloch, 1992), see "Discussion" section]. This contrasts with the abundance of results showing the crucial role of sleep

Received: 12 March 2016; Revised: 15 July 2016. Accepted: 31 July 2016

© The Author 2016. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

in promoting memory consolidation (Rasch and Born, 2013) as well as studies showing how external stimulations can improve this consolidation (Oudiette and Paller, 2013). To explain this discrepancy, the main theories on the role of sleep in memory have proposed that consolidation mechanisms directly or indirectly prevent the formation of new memories (Diekelmann and Born, 2010; Hasselmo, 1999; Tononi and Cirelli, 2014). For example, memory systems (such as hippocampal structures) could get disconnected from sensory circuits so as to prevent external input from interfering with the consolidation process (Rasch and Born, 2013). Alternatively, changes in neuromodulation occurring during sleep could impair synaptic plasticity itself and therefore the encoding of new memories (Hennevin et al., 2007; Tononi and Cirelli, 2014). Yet, recent studies have shown that even hippocampal-dependent forms of learning are possible during sleep (Arzi et al., 2012; de Lavilléon et al., 2015), putting into question the opposition between memory consolidation and memory formation.

To investigate the brain's ability to form memory traces during sleep (here, the memory of having heard a specific item), we relied on the classical old/new paradigm. For more than a century (Ebbinghaus, 1885), this approach has been used to probe recognition memory, i.e. the ability to recognize elements previously encountered. Recognition is a form of long-term and declarative memory relying on the medio-temporal lobe, which includes the hippocampus (Eichenbaum et al., 2007; Squire et al., 2007). Accordingly, participants with bilateral lesions in these areas show strong deficits in recognition memory (Reed and Squire, 1997). More precisely, the ability to recognize a previously encountered item is thought to benefit from two separate systems: an explicit episodic memory (i.e. "remembering") and an implicit sense of familiarity (i.e. "knowing") (Tulving, 1985). However, it is unclear whether these two types of information are implemented through independent neural circuits (Aggleton and Brown, 2006; Manns et al., 2003; Squire et al., 2007). Part of the difficulty in disentangling these two forms of memory stems from the difficulty to separate them at the behavioral level (Malmberg, 2008). Indeed, both implicit and explicit memories contribute to item recognition. A potential solution consists in contrasting neural correlates of recollection as a function of whether participants are aware or unaware of learning a specific content (Rosenthal et al., 2010). Sleep, in this regard, represents a unique tool to explore the formation of memories in the absence of awareness.

Here we investigated the formation of memory traces for words heard during Non-Rapid Eye-Movement (NREM) sleep. We used a paradigm in which participants were exposed to words and pseudo-words (phonologically valid but meaningless words) during wakefulness and sleep (Kouider et al., 2014). To maximize the probability that participants processed acoustic information during sleep, we asked them to perform a task on these stimuli while falling asleep. Subjects were instructed to indicate, each time a stimulus was played, whether it was a real or an invented word (lexical decision task). The task-set and stimuli presentation were held constant throughout the experiment so that participants could automatize the task while being awake and pursue it after falling asleep. Importantly, novel, unpracticed items were presented exclusively during sleep (i.e. words which were not presented during the wake session), allowing to confirm interpretations in terms of lexico-semantic processing rather than stimulus-response mapping (Abrams and Greenwald, 2000). In a previous article, we have shown the maintenance of EEG indexes of motor lateralization after stimuli onset and in accordance with the expected side of response, showing that

participants continue to classify stimuli while asleep (Kouider et al., 2014). The lexical decision task was used to prompt participants to process items at a high level of representation, as processing depth is known to influence recognition memory (Craik and Tulving, 1975). Upon awakening, participants recognized items presented during wakefulness with high accuracy. However, despite having categorized the novel items while sleeping, participants did not explicitly remember these words. Here, we show that despite the absence of explicit memory, more in-depth analyses unraveled the presence of implicit mnemonic traces both at behavioral and neural levels. Interestingly, the EEG correlates of memory for words presented during sleep markedly differed from those of the words presented in wakefulness, which were explicitly recognized. These results reveal not only that the human brain can recognize items encountered during sleep but also pinpoint critical differences in the neural implementation of explicit and implicit memories.

Material and Methods

Participants

Twenty-two (22) right-handed French speakers (16 females, age ranging from 20 to 28 years) without history of neurological or sleep disorders and with self-reported normal hearing participated in this study. Subjects had been selected based on their responses to the Epworth Sleepiness Scale (ESS) in order to target individuals who could fall asleep in a noisy, unfamiliar environment. Recruited participants had high but nonabnormal EES scores (11.95 ± 0.62 , mean \pm SEM, standard error of the mean). The day of the recordings, participants were moderately sleep deprived (30% less than their usual sleep time) and asked to avoid all exciting substances. This protocol has been approved by the local ethical committee (Conseil d'évaluation éthique pour les recherches en santé, Paris, France).

Stimuli

The auditory material consisted of 108 pairs of words and pseudo-words selected from the Lexique database (New et al., 2004). These 216 items were divided into three lists of 72 stimuli matched for their frequency, duration, and consonant-vowel structures (half CVC monosyllabic and half CV-CV disyllabic). Pseudo-words did not violate the pronunciation rules of the French language. Words were uttered by a male native French speaker and digitized at 44 100Hz. The attribution of the three lists to either the wake period, the sleep period or the new list in the old/new recognition task was counterbalanced across participants. Stimuli were presented at about 50 dB through loud speakers using the Psychtoolbox extension (Brainard, 1997) for Matlab (MathWorks Inc. Natick, MA, USA).

Experimental procedure

Participants first performed a lexical decision task on the spoken words presented every 9 s (nap session). Participants were instructed to indicate whether the spoken words existed in the French lexicon or not using response-handles placed in their left and right hands. The mapping between stimulus category (real or invented word) and the associated response side (left or right) was counterbalanced across participants. Participants initially performed this task for about 10 min with the instruction to remain awake and responsive. All the items of the wake list were played once during this initial part to ensure that all items of the wake list were processed at least once under optimal

conditions. Participants were then placed in a reclining chair; in a dark, electrically and acoustically shielded cabin. They were asked to keep their eyelids closed. Participants were authorized to fall asleep but were instructed to keep responding to auditory stimuli as long as they were awake, and to resume responding in case of an awakening.

Subjects' vigilance state was assessed online using polysomnographic and behavioral data (see below). A novel list of words and pseudo-words was presented to participants when they entered the NREM2 stages (i.e. after the first spontaneous K-complex or sleep spindle). The sleep list was played in NREM2 and NREM3 stages. However, only 11 participants entered the NREM3 stage. Participants were switched back to the wake list whenever they showed signs of arousal. This online sleep scoring was confirmed offline using standard guidelines (Iber et al., 2007). Trials associated with arousals (button presses or increase in low-amplitude fast rhythms such as alpha oscillations

or oscillations above 16 Hz for more than 3 s and stable for at least 10 s) and micro-arousals (less than 3 s) were carefully marked and corresponding items were discarded from further analyses. Indeed, our goal was here to analyze the presence of mnemonic traces to words processed during sleep and while ensuring that sleep was preserved. Details about sleep scoring and how we controlled for the presence of arousals can be found in our previous publication (Kouider et al., 2014), which includes the current dataset (see the "Sleep Lexical Decision" task). After offline confirmation of the sleep scoring, four participants could not be included in the sleep analyses (Fig. 1c) due to a low number of trials scored as sleep. One additional participant had less than 20 novel words presented during sleep and was discarded from the memory test analyses. Overall, the 17 participants included in these analyses (Figs 2–5) heard each item of the wake list 4.8 ± 0.3 times on average and each item of the sleep list 2.1 ± 0.2 times on average.

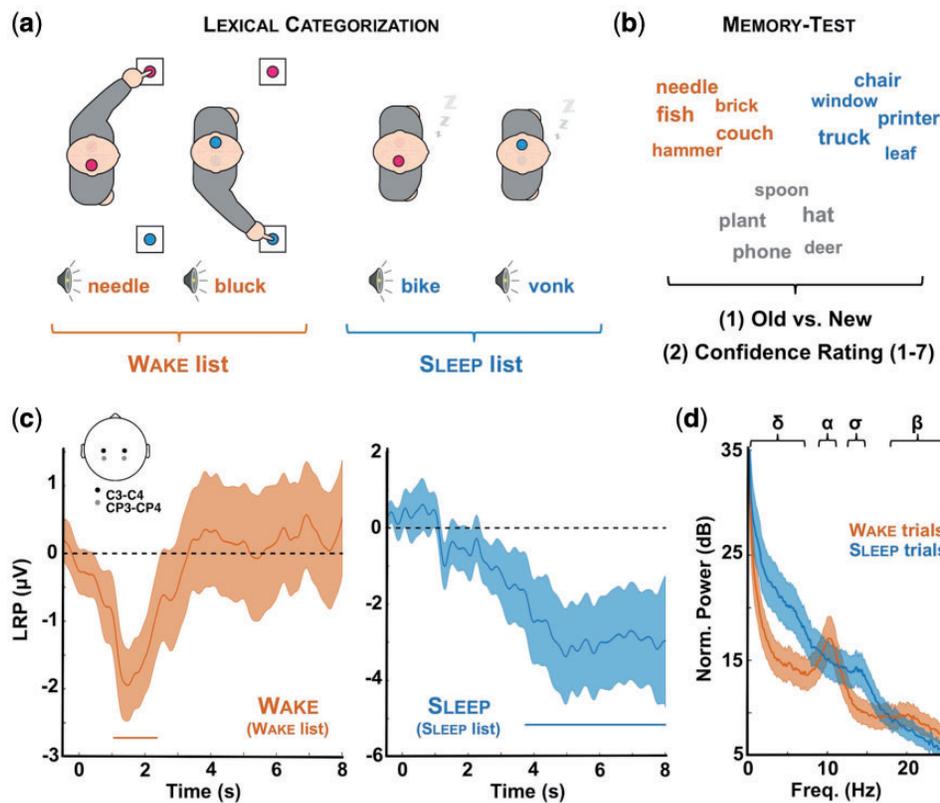


Figure 1. Experimental procedure and evidence for complex information processing during sleep. (a) Lexical categorization task: participants ($N = 22$) were asked to classify spoken words and pseudo-words while falling asleep (daytime nap). Each category (word or pseudo-word) was associated with a response side (left or right). Once asleep (see "Methods" section for criterion) and unknown to participants, novel words and pseudo-words were presented to participants (sleep list). Importantly, the initial set of stimuli (wake list) was played again to participants whenever they awoke in order to present the elements of the sleep list during sleep only. (b) Memory test: upon awakening, participants ($N = 17$) underwent a memory test during which the words presented during the nap session (red: wake list; blue: sleep list) were played along new words (gray: new list). For each word, participants were asked to indicate whether they remembered hearing the word during the nap session (old vs. new recognition test). They were then instructed to estimate their confidence in their old vs. new response using a scale from 1 (not sure at all) to 7 (absolutely sure). (c) LRPs: during the nap session, we computed LRPs during wake (red, left) and sleep (blue, right) to determine whether participants performed the lexical categorization task even when unresponsive. The LRP was computed over right and left central electrodes (see inset and "Methods" section) and averaged across participants. Shaded areas denote the standard error of the mean across participants ($N = 18$). Horizontal bars show the significant clusters ($p_{\text{cluster}} < 0.05$). The LRP determines whether neural activity is lateralized in respect to the expected side of response. A LRP was preserved in sleep, albeit delayed compared to wakefulness. (d) Average power spectra for wake (red) and sleep (blue) trials. Power spectra were computed on C3/C4 referenced to the mastoids with a FFT and averaged across subjects. Note the replacement of wake-related rhythms (α : 8–11 Hz; β : 20–30 Hz) by sleep-related oscillations (δ : 0.1–4 Hz; σ : 11–16 Hz). Shaded areas denote the standard error of the mean across participants ($N = 18$).

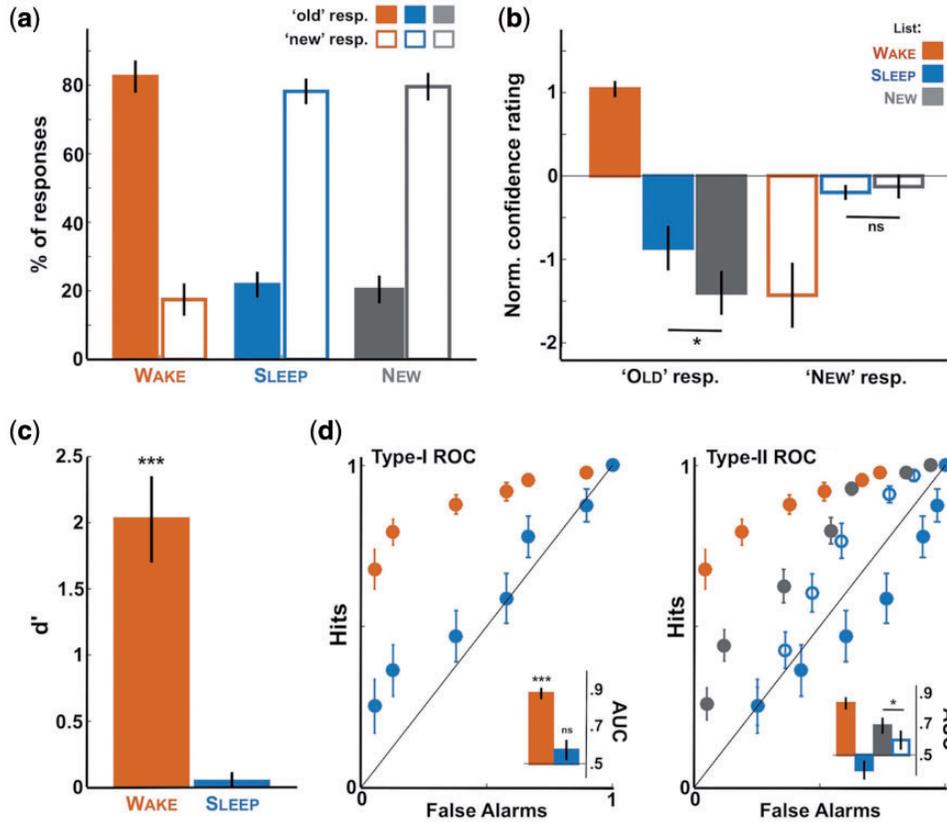


Figure 2. First- and second-order responses in the memory test. Behavioral results for the memory test. (a) Proportion of “old” and “new” responses for the wake (red), sleep (blue), and new (gray) lists (first-order responses). Proportions are averaged across participants ($N = 17$) after excluding items heard around arousals or micro-arousals for the sleep list. Note that participants responded to sleep items as if they were new. (b) Confidence rating associated with the old and new responses for the wake, sleep, and new lists (second-order responses averaged across participants). Ratings were normalized by subtracting for each participant the average confidence rating. Note the increase in confidence for sleep items in comparison to new items when participants categorized them as old. (c) Sensitivity index d' averaged across participants for the old vs. new discrimination and computed for the wake and sleep lists separately. As in panel a, the null d' reveals participants' failure to distinguish sleep and new items in their first-order responses. (d) Type-I and type-II ROC curves computed for the first order (old vs. new, left) or second order (confidence rating, right) responses. Inset shows the AUC computed for the different ROC curves and averaged across participants (0.5: area under the bisector line). In all subpanels, error-bars denote the standard error of the mean across participant. Stars atop bars denote the significance levels when comparing with 0 [t-test (u -tests for AUC values) across participants; *** $P < 0.005$, ** $P < 0.01$, * $P < 0.05$, ns: $P > 0.05$]. Stars between bars show the paired comparison between bars (paired t-tests or u -tests).

Upon awakening, after the nap session, participants were allowed a few minutes to dissipate sleep inertia and then underwent a memory test. Words (but not the pseudo-words) previously presented during the nap session were played once to participants (same voice, volume and experimental set-up) randomly intermixed with novel items (new list). Participants were first asked to indicate whether they remembered having heard the word during the nap session (“old” or “new”). They were then asked to indicate their confidence in their “old” vs. “new” response by using a scale going from 1 (“I am not sure at all”) to 7 (“I am perfectly sure”). Responses were provided with a keyboard and without time pressure. Participants were instructed to keep their eyes closed and to remain still during the presentation of the words in order to minimize movement-related artifacts in the EEG signal. The memory test was self-paced.

Behavioral analyses

Using participants' responses recorded in the memory test, we computed the average percentage of “old” and “new” responses for the different lists (first-order responses, Fig. 2a) as well as

the average confidence rating of these “old” and “new” responses (second-order responses, Fig. 2b) across participants. Confidence ratings were normalized by subtracting the average rating computed across all trials for a given participant in order to compensate for biases in the way participants scaled their own confidence. For the first-order response, we also computed a sensitivity index: the d' (Green and Swets, 1966; Macmillan, 2005). The d' provides an unbiased estimate of participants' ability to discriminate two conditions [here old words (wake or sleep list) vs. new ones]. The d' was computed as follows:

$$d' = z(\text{Hit}) - z(\text{FA}),$$

where $z(x)$ corresponds to the z -score for proportion x ; Hit corresponds to the proportion of correct responses for the words heard during the nap session (either wake or sleep list) and FA (False Alarms) corresponds to the proportions of incorrect responses for the novel words.

To assess participants' performance on the first- and second-order responses, we also computed Receiver Operating

Characteristic (ROC) curves [(Fig. 2d (Macmillan, 2005)]. Type-I ROC curves were computed for each participant using the first-order responses. For each confidence level (from 1 to 7), we computed the average proportion of hits (“old” response when an old item was presented) and of false alarms (“old” response when a novel item was presented). Figure 2d (left) shows the type-I ROC curve averaged across participants. Type-II ROC curves were computed for each participant using second-order responses (Fleming et al., 2010; Macmillan, 2005). Each confidence level was analyzed in turn (from 1 to 7): for the confidence level n , second-order hits correspond to trials in which participants scored their confidence higher or equal to n while being correct, whereas false alarms correspond to trials in which participants scored their confidence higher or equal to n while being incorrect. For each confidence level, we computed the average proportions of second-order hits and false alarms to build the type-II ROC curves. Figure 2d (right) shows the type-II ROC curve averaged across participants.

We also extracted the area under the curves (AUC) for the type-I and type-II ROC curves for each participant using the “polyarea” function in Matlab. AUC were compared with the AUC under the bisector line (0.5). Indeed, a type-I or type-II ROC curve overlapping the bisector line characterizes at-chance first- or second-order performance, respectively. Leftward deviations from the bisector line characterize above-chance performance and rightward deviations below-chance performance. One participant did not have any second-order false alarm. Its AUC was thus put to 1 (perfect performance). Excluding this participant did not change the outcome of the statistical analyses performed on type-II ROC curves.

EEG recordings

Participants were equipped for polysomnographic recordings [electroencephalography (EEG), electromyography (EMG), and electrooculography (EOG)] using a 65-channels EEG cap and additional sensors placed on participants’ skin (Electrical Geodesic Inc.). Data were acquired at 250 Hz and EEG derivations were referenced online to Cz. EOG were extracted by using electrodes placed close to the right and left canthi and referenced to the opposite mastoids. Three EMG derivations were recorded: on the chin and on the right and left abductor pollicis brevis (thumb flexor muscle) to record EMG activity associated with hand responses. EEG, EMG, and EOG data were continuously recorded in both the nap and the ensuing memory test.

Lateralized readiness potentials

The EEG data acquired during the nap has been previously analyzed. Details can be found in Kouider et al. (2014). Briefly, continuous EEG data were rereferenced to the average mastoids and high-pass filtered above 0.1 Hz (two-pass Butterworth filter at the fifth order). After a first epoching on large temporal windows centered on stimulus onset ($[-16, 16]$ s), EEG data were low-pass filtered below 30 Hz (two-pass Butterworth filter at the fifth order), epoched from -2 to 8 s and corrected for baseline activity ($[-2, 0]$ s). Trials passing an absolute threshold of $250 \mu\text{V}$ were rejected from further analyses. Lateralized readiness potentials (LRPs) were computed by subtracting the EEG signal over the right (electrodes 50 and 46 in the EGI HCGSN-64 v1 net, equivalent to C4 and CP4 in the 10/20 montage) and left (electrodes 20

and 26 in the EGI HCGSN-64 v1 net, equivalent to C3 and CP3 in the 10/20 montage) electrodes:

$$\text{LRP} = \frac{(C3/CP3_{\text{right-hand}} - C3/CP3_{\text{left-hand}}) + (C4/CP4_{\text{left-hand}} - C4/CP4_{\text{right-hand}})}{2}.$$

LRP quantifies the lateralization of brain activity toward the expected side of response (Smulders et al., 2012). Since novel words were presented during sleep, and since the decision to prepare for the right or left responses was based on stimulus category, the observation of an LRP also implies that auditory information was processed at a high (i.e. lexico-semantic) level of representation.

We also computed the spectral power around stimulus presentation (-2 to 6 s around stimulus onset) in order to compare the neural dynamics in wake and sleep trials. Power spectra were computed on C3–C4 electrodes over each epoch, using a fast Fourier transform (FFT). For each epoch, power was normalized by the power within higher frequencies (35–45 Hz) and expressed in decibels. The power spectra averaged across participants for sleep and wake trials show clear differences (Fig. 3d) with, notably, the replacement of wake rhythms (α : 8–11 Hz; β : 20–30 Hz) by sleep oscillations (θ : 0.1–4 Hz; σ : 11–16 Hz).

Event-related potentials

Upon awakening, participants underwent a memory test during which EEG data were recorded along participants’ behavioral responses. We computed event-related potentials (ERPs) time-locked to stimuli onset for the three lists (wake, sleep, and novel words). EEG data were rereferenced to the averaged mastoids and high-pass filtered above 0.1 Hz (two-pass Butterworth filter at the fifth order). The EEG signal was then epoched on temporal windows time-locked to stimulus onset ($[-4, 4]$ s) and low-pass filtered below 30 Hz (two-pass Butterworth filter at the fifth order). Next, EEG data were epoched from -0.2 to 1.5 s around stimulus onset and corrected for baseline activity ($[-0.2, 0]$ s). Finally, EEG data were de-noised using the joint decorrelation approach by optimizing the repeatability across all trials (de Cheveigné and Parra, 2014). Briefly, a principal component analysis (PCA) was applied to the average ERP computed across all trials and for a given participant (i.e. regardless of the stimulus list). Components were sorted according to their participation to the average ERP. The first 10 components, characterized by the strongest mean effect relative to overall variability, were used as a bias filter on the single-trial EEG data. EEG data were then averaged for all the trials of a given list and for each participant. Figure 3 shows the corresponding ERP traces averaged across participants. Differences between ERP traces could be observed when comparing the wake and new lists (P3 and Late Negativity, Fig. 3) or the sleep and new lists (Centro-Parietal Negativity, Fig. 3).

Decoding

For each participant and EEG sensor, we extracted the signal averaged over the “P3” cluster (wake vs. new comparison) for each and every trial. The cluster was defined as a $[0.73, 0.86]$ ms window post-stimulus onset on parieto-occipital electrodes (electrodes showing the P3 effect: $P < 0.05$ when comparing the wake and new lists across participants). A leave-one-out approach was performed at the subject level. For all subjects but one, the cluster data were z-scored for each channel and trials,

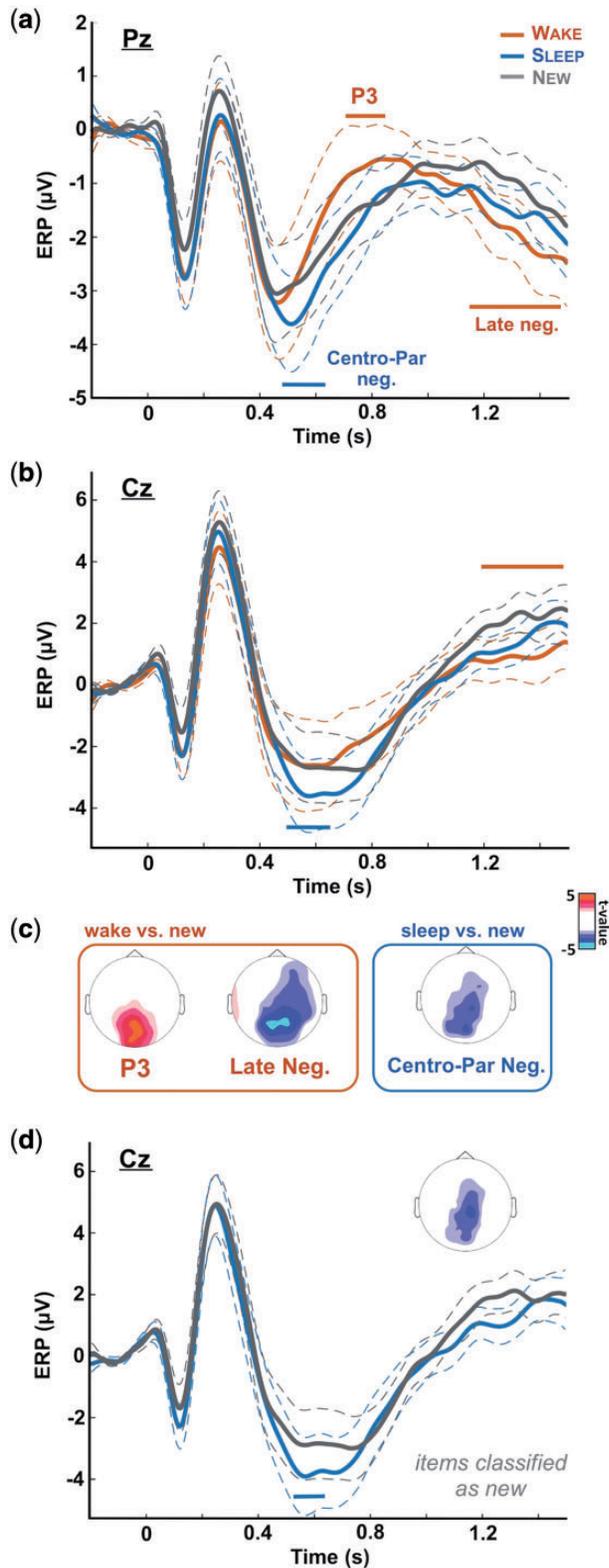


Figure 3. ERP to wake, sleep, and new words. The EEG signal time-locked to sound onset was averaged across participants ($N = 17$) for the three lists separately (red: wake; blue: sleep; gray: new lists). (a) Shows the ERPs for electrode Pz and (b) for electrode Cz. Horizontal bars show the significant clusters for the comparison between the wake list (red bars) or the sleep list (blue bars) and the new list ($p_{\text{cluster}} < 0.05$). The scalp topographies corresponding to the different

and aggregated across participants. The corresponding values (training set) were then used to fit a linear regression classifier that was then applied to the remaining participant (test set) in order to predict trials' category (wake or new). This procedure was iterated until all subjects had been included in the test set. The predictions were compared to the actual categories and both d' and AUC were computed for each participant (Fig. 4). The same procedure was applied to the sleep vs. new cluster (Fig. 3, [0.49, 0.63] ms post-stimulus onset over centro-parietal electrodes).

Time-frequency decomposition

We computed the time-frequency decomposition of the EEG signal in response to sounds. To do so, a FFT was applied to the

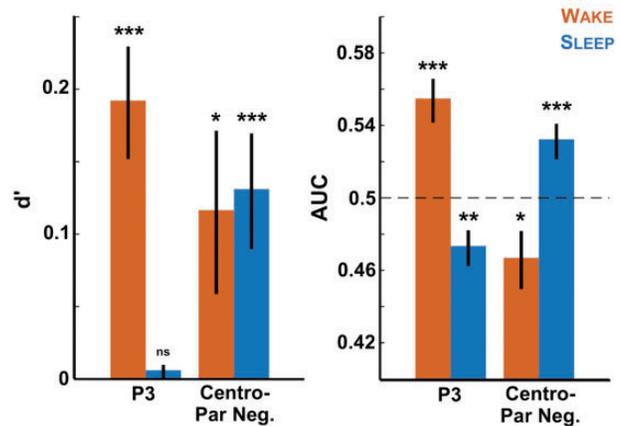


Figure 4. Decoding stimulus category with the EEG signal. The EEG amplitude over the clusters observed in Fig. 3 ('P3' cluster: [0.73, 0.86] s over parieto-occipital electrodes; "Centro-Parietal Negativity" cluster: [0.49, 0.62] s over centro-parietal electrodes) was retrieved on a single trial basis for all participants. A leave-one-out approach was applied at the subject-level to fit a logistic regression (see "Methods" section), which was then used to predict stimulus category (sleep vs. new or wake vs. new). The sensitivity index d' (left) and AUC (right) of this prediction was computed for each participant and each contrast and averaged across participants. The bars show therefore the extent to which we can predict if an item (red: wake item; blue: sleep item) had been previously heard from the EEG signal. Stars atop bars show the Monte Carlo P -values comparing the real decoding performance with surrogate datasets obtained through the permutation of trial categories (see "Methods" section: *** $P < 0.001$, ** $P < 0.005$, * $P < 0.05$, ns: $P > 0.05$). Note that the P3 cluster allows classifying only wake items while the negative sleep vs. new cluster could differentiate both wake and sleep items from new ones.

clusters are shown on (c) [t -values computed by comparing the signal averaged over the cluster time-window between the wake (or sleep) and new lists, nonsignificant t -values ($P > 0.05$) were set to 0]. Note the classical posterior positivity when comparing wake (explicitly remembered) and new items ('P3'). For sleep items, an opposite negativity (centro-parietal negativity) was observed in lieu of a positivity, despite the absence of explicit recognition at the behavioral level (Fig. 2). (d) Shows the same ERPs as in (b) (Cz electrode) for sleep and new items declared as new. The inset shows the topography of the significant cluster observed between the sleep and new lists. In (a), (b), and (d), dotted lines show the standard error of the mean computed across participants ($N = 17$).

EEG signal on band-passed de-noised stimulus-locked data ([0.1, 40] Hz, [-2, 2] s) using a window of 800 ms. The average power was computed across all trials (Fig. 5 top) or across all trials for a given list (Fig. 5 bottom) and expressed as the log ratio of the power at a given time and frequency over the power average over the baseline for the corresponding frequency. We later focused on the alpha band ([8, 12] Hz) by averaging the power within the corresponding frequency range.

Statistics

Parametric statistics were used here (Student *t*-tests to compare conditions or a condition with chance-level, ANOVA for analyses of variance) when the corresponding data could be considered normally distributed. Nonparametric tests (e.g.

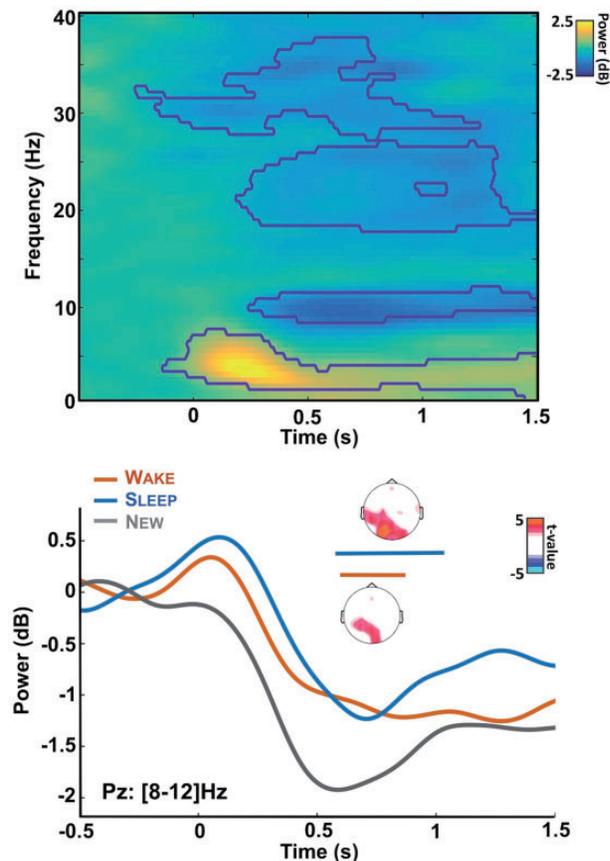


Figure 5. Alpha desynchronization during memory recognition. Top: time-frequency decomposition of the EEG signal recorded at electrode Pz and for all trials. Power is expressed as a log ratio of the baseline activity and was averaged across participants ($N = 17$). The contours show the significant clusters ($p_{\text{cluster}} < 0.05$). After stimuli onset, clear modulations of the power can be observed in the theta ([3–8] Hz), alpha ([8–12] Hz) and beta (>20 Hz) bands. Bottom: Power modulation for the alpha band ([8–12] Hz) over electrode Pz for the different lists of stimuli (red: wake, blue: sleep, gray: new list). Horizontal bars show the significant cluster when comparing the wake and new lists (red: $p_{\text{cluster}} = 0.02$) or the sleep and new lists (blue: $p_{\text{cluster}} = 0.009$) across participants. Insets show the corresponding topographies for these clusters: power was averaged on the clusters' time-window and *t*-values were computed by comparing the corresponding power for the wake (or sleep) list and the new one. Nonsignificant *t*-values ($P > 0.05$, uncorrected) were set to 0.

Mann–Whitney *U* test to compared two conditions) were used otherwise.

To correct for multiple comparisons in time-plots and time-frequency-plots, we used a principled approach called “cluster permutation” (Maris and Oostenveld, 2007). Each cluster was constituted by the samples [in one dimension (time) or two dimensions (time and frequency)] that consecutively passed a specified threshold (for time-plots: $P < 0.1$; for time-frequency-plots: $P < 0.01$). It has been shown that the rate of type-I errors (false positives) was immune to the choice of this cluster-defining threshold (Maris and Oostenveld, 2007). The cluster statistics were chosen as the sum of the *t*-values of all the samples within the cluster. Then, we compared the cluster statistics of each cluster with the maximum cluster statistics of 1000 random permutations (Monte Carlo method). From this comparison, we obtained a Monte Carlo *P* value: the cluster *P* value (P_{cluster}).

We also used the EEG signal to try to predict stimulus category (see above). To determine whether the accuracies of such decoding were above chance-level, AUC and d' values were compared to AUC and d' values computed on surrogate datasets in which trial conditions of the training set were shuffled within each participant ($N = 1000$ permutations). The position of the real AUC (or d') value within the surrogate values' distribution was used to compute a Monte Carlo *P* value reported in the “Results” section.

When examining the presence of mnemonic traces in the sleep lists, we obtained several null results. To check the informativeness of these null results, we computed Bayes Factors given the effect observed in wakefulness (Dienes, 2014). Bayes factors allowed assessing whether a null result is in favor of the null hypothesis or reflects data's lack of sensitivity. The larger the Bayes Factor, the more the data is in favor of the null hypothesis.

Results

In a recent publication, we have shown that sleepers can maintain complex and flexible processing of sensory information during NREM sleep (Kouider et al., 2014). Such feat was not accompanied by any explicit memory, contrasting with what happens when we are awake. Here we investigated the presence, in the behavioral and EEG data, of implicit mnemonic traces for words heard during sleep.

Memory traces for words heard while awake

Upon awakening, in the memory test, we used both first- and second-order responses to determine the presence of a mnemonic trace. We first checked whether the list category (wake, sleep, or new) and the list itself (i.e. which list was defined as the wake, sleep, or new list for a given participant) had an effect on accuracy (first-order response) and confidence rating (second-order response). Using ANOVAs, we observed, for both correctness and confidence, an effect of the list category [$F(2) = 80.5$ and $F(2) = 5.67$, $P = 4.10^{-15}$ and $P = 0.007$ resp.] but not of the list itself [$F(2) = 2.39$ and $F(2) = 0.03$, $P = 0.10$ and $P = 0.97$ resp.]. There was no significant interaction [$F(4) = 0.39$ and $F(4) = 0.60$, $P = 0.82$ and $P = 0.66$ resp.], suggesting that the different lists were correctly balanced between vigilance states. Thus, a difference in either first- or second-order responses between the wake (or sleep) list and the new one can be interpreted as evidence for the existence of a mnemonic trace. In addition and following previous work, first-order responses

were used to determine the explicit nature of a memory (Kouider and Dehaene, 2007), that is to say that words leading to above-chance first-order performance were interpreted as being explicitly remembered. On the contrary, words showing significant effect without above-chance first-order performance were interpreted as being implicitly recognized (Chong et al., 2014; Rosenthal et al., 2016, 2010).

Words presented during wakefulness were explicitly recognized as “old” words with a high degree of accuracy ($83 \pm 4\%$, Fig. 2a). Such performance led to a high d' index when contrasting these wake items with novel ones [t-test comparing with 0, $t(16) = 6.62$, $P = 6.10^{-6}$]. Participants also attributed high levels of confidence when they were correctly recognizing an item (6.2 ± 0.17 over 7) and low levels of confidence when missing to identify an item previously heard (3.7 ± 0.40 over 7, see Fig. 2b for normalized values). Such pattern of results led to type-I and type-II ROC curves clearly distinct from the bisector line. Accordingly, the wake list AUCs were highly significant when compared to bisector's AUC (u -test: $P < 5.10^{-4}$ for both the type-I and type-II ROC curves). The type-I ROC curve in particular had an asymmetric shape that is typical of strong and explicit memory traces (Squire et al., 2007). Overall, participants were unsurprisingly able to accurately and confidently discriminate items heard while awake from new ones.

These behavioral effects were accompanied by differences in the ERP between the wake and new lists (Fig. 3). Namely, a positivity over occipito-parietal electrodes was observed around 700 ms when comparing the wake and new lists across participant (cluster on Pz: [0.72, 0.86] s, $p_{\text{cluster}} = 0.026$) in accordance with previously published results (Kayser et al., 2007; Rugg et al., 1998; Voss and Paller, 2007). This rather delayed difference affects the third positivity within the stimulus-locked ERPs and was thus termed the “P3 effect” in the literature. Such late (~700 ms post-stimulus) memory-related effect should not be confounded with the earlier P300, which indexes attention and expectation during perception (Polich, 2007). Another negativity, maximal over parietal electrodes, was also observed later in time (cluster on Pz: [1.17, 1.40] s, $p_{\text{cluster}} = 0.003$) and will be referred here as the “late negativity” according to previously published work (Kayser et al., 2007).

Implicit memory traces for words heard during sleep in the absence of explicit recognition

What about sleep? First of all, and as previously reported (Kouider et al., 2014), participants did not explicitly recognize the items presented during NREM sleep despite having processed them at a high level of representation (Fig. 1c). Indeed, the pattern of “old/new” responses for the sleep list was strikingly similar to the new list (Fig. 2a) leading to null d' [t-test comparison to 0: $t(16) = 0.73$, $P = 0.25$]. To determine whether this null result was informative and not reflecting data insensitivity, we computed the Bayes Factor associated with the effect on the sleep list (on d') when considering the effect on the wake list (Dienes, 2014). A Bayes Factor superior to 1000 indicates very strong evidence for the null hypothesis (Kass and Raftery, 1995). In line with the null d' for the sleep list, the type-I ROC curve for sleep items did not deviate from the bisector line (Fig. 2d, left; u -test comparison to 0.5: $t(16) = 1.34$, $P = 0.18$; Bayes factor >1000).

For second-order responses, we observed that although participants sometimes responded “old” for both sleep and novel items, they tended to give low-confidence ratings in such cases (Fig. 2b). Such low-confidence ratings reflect the absence of

explicit memory for the sleep list. Importantly however, we found that the sleep and new lists did not lead to identical second-order responses (Fig. 2b) contrary to first-order responses (Fig. 2a). Indeed, participants were more confident when responding “old” for sleep items compared to genuinely novel word [paired t-test: $t(15) = 2.3$, $P = 0.036$]. The fact that participants rated their responses differently for the sleep items compared to new ones (second-order responses) without, crucially, being able to explicitly (first-order responses) differentiate these words, advocates for the implicit nature of such memory trace. Thus, our results suggest the presence of an implicit mnemonic trace for words heard during sleep.

This effect of sleep exposure on confidence was confirmed when examining the type-II ROC curves for the sleep and novel lists (Fig. 2d). First of all, the corresponding curves drastically differed (the wake list is below the bisector line, the sleep list above). But, this difference is due to the way these curves are computed and the fact that an “old” response is considered as being correct for the sleep list and “incorrect” for the new list. Nonetheless, when recomputing the sleep-list ROC curves while considering the “old” response as being incorrect (i.e. when considering that participants did not have any explicit recollection of the sleep items), such sleep (unfilled blue dots) and new lists type-II ROC curves still differed, with the sleep list showing a smaller AUC [paired u -test comparing AUCs: $t(16) = -2.30$, $P = 0.022$]. This result provides further evidence that sleep items were not merely processed as novel items.

Thus, the presence of a second-order effect in the absence of first-order difference suggests the existence of an implicit mnemonic trace for items heard during sleep. The pattern observed in the sleep-list type-II ROC curve can be interpreted as a conflict between the absence of explicit memory for sleep words and the presence of an implicit mnemonic trace (see “Discussion” section).

EEG evidence for implicit memory traces for words heard during sleep

Differences between the sleep list and the novel list were also observed when examining the corresponding ERPs (Fig. 3). A “centro-parietal negativity” was observed when comparing the sleep and novel lists (cluster on Pz: [0.48, 0.62] s, $P = 0.006$; on Cz: [0.49 0.63]s, $P = 0.044$), suggesting that the two lists were not processed identically. In addition, this centro-parietal negativity was present even when focusing on the stimuli categorized as “new” (Fig. 3d: [0.53, 0.63]s, $p_{\text{cluster}} = 0.043$). On the other hand, there was no P3 or Late Negativity difference when contrasting the sleep list with the new list. This absence of the neural signatures of explicit recognition again suggests that the sleep list was not processed as the wake list. The centro-parietal negativity observed for the sleep lists actually overlapped spatially and temporally with the P3 positivity but with an opposite sign. Finally, this centro-parietal negativity occurred long before participants' motor responses (first-order responses: 1.94 ± 0.08 s; second-order responses: 3.39 ± 0.18 s; mean post-stimulus onset \pm SEM across 17 participants) making it unlikely that the observed potential is due to a motor component.

To verify that the clusters reported above were not impacted by the preprocessing and in particular by the data-dependent joint-decorrelation procedure (see “Methods” section), we created surrogate datasets ($N = 1000$ for the 17 participants) in which the very same preprocessing steps were applied but experimental conditions were shuffled across trials. The

differences, obtained in the original dataset, between the ERPs of the wake list or the sleep list and the ERPs of the new list were then compared to the same differences obtained in the surrogate datasets. Monte Carlo *P*-values were computed to estimate the likelihood that the differences observed in the real dataset were significant. This procedure led to similar results as reported above for both the wake list and sleep list clusters (e.g. P3 effect on Pz: [0.72, 0.86] s, $p_{\text{cluster}} = 0.017$; Late-Negativity on Pz: [1.17, 1.40] s, $p_{\text{cluster}} < 10^{-4}$).

We further checked whether the differences observed in the ERP waveforms could allow us to predict stimuli category. To do so, we extracted the average voltage over the P3 cluster and the “centro-parietal negativity” for the different stimuli lists. For each participant, we trained and tested a classifier using a leave-one-out approach (see “Methods” section) to predict stimuli category based on the EEG signal. Figure 4 shows that such procedure led to above-chance levels discrimination for the wake vs. new lists discrimination when using the P3 cluster (both *d'* and AUC: Monte Carlo $P < 0.001$, see “Methods” section). Importantly, when using the EEG data computed over the centro-parietal cluster associated with the sleep list, we could predict the stimuli category better than chance when comparing the sleep and the new lists (both *d'* and AUC: Monte Carlo $P < 0.001$, see “Methods” section). In addition, we could also predict stimuli category for the sleep vs. new contrast when using the P3 cluster (AUC: Monte Carlo $P < 0.005$) and for the wake vs. new contrast when focusing on the centro-parietal negativity (AUC: Monte Carlo $P < 0.05$). Note that differences between the significance level of the *d'* and AUC values stress the increase sensitivity of AUC computation in assessing decoding performance.

Finally, we examined the time-frequency decomposition of the EEG signal in response to the items of the different lists (Fig. 5). When pooling all words from all lists, it appears that stimuli presentation modulates the EEG signal in three distinct frequency bands. Stimulus onset was followed by a synchronization within the theta band (i.e. increase in power for the [4, 8] Hz band) and a desynchronization (decrease in power) within the alpha ([8, 12] Hz) and beta band (>16 Hz). Interestingly, variation in the theta and alpha power has been associated with changes in recollection performance (Klimesch, 1999; Klimesch et al., 1997). We did not observe any difference between lists in the theta band. However, the new list elicited stronger alpha desynchronization over occipital regions compared to the wake list ($p_{\text{cluster}} = 0.02$, Fig. 5 bottom) in accordance with previous findings (Klimesch et al., 1997). Importantly, a similar and even stronger effect was observed when contrasting the wake and sleep lists ($p_{\text{cluster}} = 0.009$). Once again, the brain response to the sleep and new lists show that despite equivalent first-order responses, the sleep items were not processed as novel words.

Overall, both behavioral and EEG data reveal the presence of implicit mnemonic traces for words heard during sleep. However, the opposite effects observed in the ERP between the wake and sleep lists suggest that the memory traces associated with wake and sleep items are qualitatively distinct.

Discussion

Memory for words heard during sleep

Until recently, scientific efforts to understand whether humans can learn while sleeping had remained largely inconclusive due to the paucity of positive results and doubts raised by methodological flaws (Bruce et al., 1970; Webb, 1990; Wood, 1990). At first

sight, our results are consistent with previous studies showing no memory for the words presented during sleep (Cox et al., 2014; Emmons and Simon, 1956; Wood et al., 1992). Indeed, while words presented during wakefulness elicited close-to-perfect explicit recognition (Fig. 2) and classical ERP signatures of recollection (Fig. 3, P3 effect), items presented during sleep seemed to be processed as new items when considering these two markers (Figs 2 and 4), evidencing the absence of explicit recognition.

However, a more detailed investigation of both behavioral and EEG data revealed differences between the sleep and new lists, suggesting the presence of subtler mnemonic traces. Sleep items that were declared as previously encountered elicited higher confidence judgments than sleep items judged as “old” (Fig. 2b). This difference in confidence estimation between the wake and sleep lists was confirmed when analyzing the type-II ROC curves (Fig. 2d). Analyzing brain responses to the sleep and new lists confirmed that these two lists were not processed identically (Fig. 3), with sleep items eliciting a larger centro-parietal negativity around 500 ms. This centro-parietal negativity was present even when restricting our analysis to items categorized as new by the participants (Fig. 3d). When extracting the EEG signal over the corresponding cluster, a classifier could separate novel from sleep items with above-chance performance (Fig. 4), which participants themselves could not do. Differences between the sleep and new lists could also be evidenced when examining the stimulus-related alpha desynchronization (Fig. 5). The EEG signal contains therefore information about the fact that sleep items had been previously heard but participants did not use this information to perform the old/new task.

Nature of the memory traces formed in sleep

Although behavioral and EEG data indicate that words presented during sleep left a trace in participants' brain, such traces seem quite different from the explicit memory formed during wakefulness. Indeed, wake items were explicitly recognized (high first- and second-order performance, Fig. 2). On the contrary, sleep items did not lead to above chance first-order performance. The memory effect was restricted, at the behavioral level, to second-order responses. An improvement of second-order responses with at-chance first-order responses has been often interpreted as the manifestation of an unconscious (i.e. implicit) memory trace (Jachs et al., 2015; Rosenthal et al., 2016, 2010; Scott et al., 2014; Tulving, 1985).

In fact, the patterns of results observed on confidence rating (increased confidence for sleep items correctly categorized as old compared to new one but lower confidence ratings compared to sleep items incorrectly categorized as new, see Fig. 2b) could be interpreted as a conflict between the absence of explicit memory for sleep items and the presence of implicit mnemonic traces. Under such conditions, when the implicit memory trace is weak or absent, participants would tend to process sleep items as new with a rather high degree of confidence. On the contrary, when the implicit trace is stronger, participants would declare such items as old but with a rather low degree of confidence due to the lack of explicit recognition. Nonetheless, these items would elicit higher confidence ratings compared to incorrectly classified new items for which no implicit memory trace exists. As a result, second-order responses exhibit a complex pattern, whereby type-II ROC curves indicate a trend for below-chance second-order performance (Fig. 2d).

Thus, our results could fit with the classical distinction between “knowing” (explicit recollection) and “remembering” (implicit recognition) (Tulving, 1985). According to this view, the explicit recollection and familiarity both contribute to recognition but recruit different memory systems (Rugg *et al.*, 1998; Vilberg and Rugg, 2008; Voss and Paller, 2007). However, disentangling the contribution of explicit and implicit memory at the behavioral level and in their neural signature is very complex (Malmberg, 2008; Rotello *et al.*, 2004; Voss and Paller, 2007). The pattern of behavioral results observed here does not indeed constitute irrefutable evidence for the implicit nature of the mnemonic traces formed during sleep. Further investigations will be needed to definitely clarify this point but the fact that participants were in a state of minimal, if not absent, consciousness during the encoding period [i.e. NREM sleep (Nir *et al.*, 2013)] argues in favor of an implicit determinant of learning. As a consequence and contrary to many studies in which the explicit/implicit nature of a stimulus is manipulated by processing depth [e.g. (Rugg *et al.*, 1998)], we can here better isolate the neural activity underlying implicit memory. Interestingly, many studies showed quantitative rather than qualitative changes between the neural correlates of implicit and explicit memory (Allan *et al.*, 1998). Here, on the contrary, sleep items elicited a centro-parietal negativity that appeared as the opposite of the classical signature of explicit recollection (i.e. P3 effect; see Fig. 3).

We further checked that this centro-parietal negativity depended on the stimuli and not subjects’ response by focusing on sleep and new items both categorized as “new” by participants. In such case, behavioral responses were identical for the two conditions. Nevertheless, a similar cluster was observed over Pz ($[0.48, 0.62]$ s, $p_{\text{cluster}} = 0.01$), suggesting that the centro-parietal negativity depends on prior exposure and not on participants’ choice.

Alternatively, it has been proposed that the difference between explicit and implicit memory could stem from a difference in memory strength rather than different neural sources (Squire *et al.*, 2007). However, such view would predict an ERP modulation similar to the wake list, albeit weaker, for the sleep list in lieu of the opposite effects observed. While both familiarity and recollection could be encoded in similar structures within the medio-temporal lobe (Squire *et al.*, 2007), our results suggest that they can be dissociated under certain conditions.

Learning or consolidating during sleep?

Sleep has often been seen as a state promoting memory consolidation to the detriment of memory encoding (Hasselmo, 1999; Hennevin *et al.*, 2007; Tononi and Cirelli, 2014). It has been proposed that the low-level of acetylcholine in NREM sleep as well as the relative disconnection of the sleepers from its environment would impair the formation of new memories. However, recent research showed that the role of sleep in memory consolidation does not necessarily preclude any sleep-learning. Accordingly, animals and humans can be conditioned during sleep (Arzi *et al.*, 2014, 2012; de Lavilléon *et al.*, 2015; Hennevin *et al.*, 2007, 1995) even for hippocampus-dependent forms of learning. One key element potentially explaining the success of these studies consists in their ability to bypass sensory isolation by providing either olfactory information (Arzi *et al.*, 2012), which does not transit by thalamic relays (Jones, 2007), or by using intracranial stimulations (de Lavilléon *et al.*, 2015). Here, we similarly ensure that the stimuli were processed at a high level of representation by checking the preservation of task-related

motor preparation indexes (Fig. 1c) that can be observed if and only if the novel information is correctly processed (Kouider *et al.*, 2014). It is also important to note that we studied mostly light stages of NREM sleep (nap studies). It is possible that light NREM is a stage more favorable to the processing of external information and the formation of memory in opposition to deeper stages of sleep in which sensory isolation and changes in neuro-modulation are more pronounced (Genzel *et al.*, 2014).

Conclusions

The formation of memory during sleep is thus possible but is a highly constrained phenomenon with little effect at the behavioral level. Nonetheless, sleep-learning resulted here in distinctive neural responses upon awakening, demonstrating the recruitment of implicit learning mechanisms during NREM sleep. Further research is needed to elucidate whether sleep-learning can be optimized when taking into consideration sleep stages or sleep rhythms, as it is the case for memory consolidation (Batterink *et al.*, 2016).

Acknowledgements

We thank A. de Cheveigné, D. Arzounian, and J. Sackur for their help regarding data analysis. Data is available on request.

Conflicts of interest statement: None declared.

Funding

This research was supported by ANR grants (ANR-10-LABX-0087 and ANR-10-IDEX-0001-02), by the European Research Council (ERC project METAWARE to S.K.) and by the Ministère de la Recherche and the Société Française de Recherche et Médecine du Sommeil (T.A.).

References

- Abrams RL, Greenwald AG. Parts outweigh the whole (word) in unconscious analysis of meaning. *Psychol Sci* 2000;11:118–24.
- Aggleton JP, Brown MW. Interleaving brain systems for episodic and recognition memory. *Trends Cogn Sci* 2006;10:455–63.
- Allan K, Wilding EL, Rugg MD. Electrophysiological evidence for dissociable processes contributing to recollection. *Acta Psychol (Amst)* 1998;98:231–52.
- Andrillon T, Kouider S, Agus TR, *et al.* Perceptual learning of acoustic noise generates memory-evoked potentials. *Curr Biol* 2015;25:2823–29.
- Arzi A, Holtzman Y, Samnon P, *et al.* Olfactory aversive conditioning during sleep reduces cigarette-smoking behavior. *J Neurosci* 2014;34:15382–93.
- Arzi A, Shedlesky L, Ben-Shaul M, *et al.* Humans can learn new information during sleep. *Nat Neurosci* 2012;15:1460–5.
- Bastuji H, Perrin F, Garcia-Larrea L. Semantic analysis of auditory input during sleep: studies with event related potentials. *Int J Psychophysiol* 2002;46:243–55.
- Batterink LJ, Creery JD, Paller KA. Phase of spontaneous slow oscillations during sleep influences memory-related processing of auditory cues. *J Neurosci* 2016;36:1401–9.
- Brainard DH. The psychophysics toolbox. *Spat Vis* 1997;10:433–6.
- Brualla J, Romero MF, Serrano M, *et al.* Auditory event-related potentials to semantic priming during sleep. *Electroenc Clin Neurophysiol* 1998;108:283–90.

- Bruce DJ, Evans CR, Fenwick PBC, et al. Effect of presenting novel verbal material during slow-wave sleep. *Nature* 1970;225:873–4.
- Chong TTJ, Husain M, Rosenthal CR. Recognizing the unconscious. *Curr Biol* 2014;24:R1033–5. doi:10.1016/j.cub.2014.09.035
- Cox R, Korjoukov I, de Boer M, et al. Sound asleep: processing and retention of slow oscillation phase-targeted stimuli. *PLoS ONE* 2014;9:e101567.
- Craik FIM, Tulving E. Depth of processing and the retention of words in episodic memory. *J Exp Psychol Gen* 1975;104:268–94.
- Czisch M, Wehrle R, Stiegler A, et al. Acoustic oddball during NREM sleep: a combined EEG/fMRI Study. *PLoS ONE* 2009;4:e6749. doi:10.1371/journal.pone.0006749
- de Cheveigné A, Parra LC. Joint decorrelation, a versatile tool for multichannel data analysis. *NeuroImage* 2014;98:487–505.
- de Lavilléon G, Lacroix MM, Rondi-Reig L, et al. Explicit memory creation during sleep demonstrates a causal role of place cells in navigation. *Nat Neurosci* 2015;18:493–5.
- Diekelmann S, Born J. The memory function of sleep. *Nat Rev Neurosci* 2010;11:114–26.
- Dienes Z. Using Bayes to get the most out of non-significant results. *Front Psychol* 2014;5:781.
- Ebbinghaus H. 1885. *Memory: A Contribution to Experimental Psychology*. New York City, NY: Teachers College, Columbia University.
- Eichenbaum H, Yonelinas AP, Ranganath C. The medial temporal lobe and recognition memory. *Annu Rev Neurosci* 2007;30:123–52.
- Emmons WH, Simon CW. The non-recall of material presented during sleep. *Am J Psychol* 1956;69:76–81.
- Fleming SM, Weil RS, Nagy Z, et al. Relating introspective accuracy to individual differences in brain structure. *Science* 2010;329:1541–3.
- Genzel L, Kroes MCW, Dresler M, et al. Light sleep versus slow wave sleep in memory consolidation: a question of global versus local processes? *Trends Neurosci* 2014;37:10–9.
- Green DM, Swets JA. 1966. *Signal Detection Theory and Psychophysics*. Los Altos Hills, CA: Peninsula Publishing.
- Hasselmo ME. Neuromodulation: acetylcholine and memory consolidation. *Trends Cogn Sci* 1999;3:351–9.
- Hennevin E, Hars B, Maho C, et al. Processing of learned information in paradoxical sleep: relevance for memory. *Behav Brain Res* 1995;69:125–35.
- Hennevin E, Huetz C, Edeline JM. Neural representations during sleep: from sensory processing to memory traces. *Neurobiol Learn Mem* 2007;87:416–40.
- Ibanez A, Lopez V, Cornejo C. ERPs and contextual semantic discrimination: degrees of congruence in wakefulness and sleep. *Brain Lang* 2006;98:264–75.
- Iber C, Ancoli-Israel S, Chesson A, et al. 2007. *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*. Westchester, IL: American Academy of Sleep Medicine.
- Ikeda K, Morotomi T. Classical conditioning during human NREM sleep and response transfer to wakefulness. *Sleep* 1996;19:72–74.
- Issa EB, Wang X. Sensory responses during sleep in primate primary and secondary auditory cortex. *J Neurosci* 2008;28:14467–80.
- Jachs B, Blanco MJ, Grantham-Hill S, et al. On the independence of visual awareness and metacognition: a signal detection theoretic analysis. *J Exp Psychol Hum Percept Perform* 2015;41:269–76.
- Jones EG. 2007. *The Thalamus*, 2nd edn. Cambridge, UK; New York: Cambridge University Press.
- Kass RE, Raftery AE. Bayes Factors. *J Am Stat Assoc* 1995;90:773–95.
- Kayser J, Tenke CE, Gates NA, et al. Reference-independent ERP old/new effects of auditory and visual word recognition memory: joint extraction of stimulus- and response-locked neuronal generator patterns. *Psychophysiology* 2007;44:949–67.
- Klimesch W. EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res Brain Res Rev* 1999;29:169–95.
- Klimesch W, Doppelmayr M, Schimke H, et al. Theta synchronization and alpha desynchronization in a memory task. *Psychophysiology* 1997;34:169–76.
- Kolb B, Whishaw IQ. Brain plasticity and behavior. *Annu Rev Psychol* 1998;49:43–64.
- Kouider S, Andrillon T, Barbosa LS, et al. Inducing task-relevant responses to speech in the sleeping brain. *Curr Biol* 2014;24:2208–14. doi:10.1016/j.cub.2014.08.016
- Kouider S, Dehaene S. Levels of processing during non-conscious perception: a critical review of visual masking. *Philos Trans R Soc Lond B Biol Sci* 2007;362:857–75.
- Macmillan NA. 2005. *Detection Theory: A User's Guide*, 2nd edn. Mahwah, NJ: Lawrence Erlbaum Associates.
- Maho C, Bloch V. Responses of hippocampal cells can be conditioned during paradoxical sleep. *Brain Res* 1992;581:115–22.
- Malmberg KJ. Recognition memory: a review of the critical findings and an integrated theory for relating them. *Cognit Psychol* 2008;57:335–84.
- Manns JR, Hopkins RO, Reed JM, et al. Recognition memory and the human Hippocampus. *Neuron* 2003;37:171–80.
- Maris E, Oostenveld R. Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 2007;164:177–90.
- New B, Pallier C, Brysbaert M, et al. Lexique 2: a new French lexical database. *Behav Res Methods Instrum Comput J Psychon Soc Inc* 2004;36:516–24.
- Nir Y, Massimini M, Boly M, et al. Sleep and consciousness. In: Cavanna AE, Nani A, Blumenfeld H, Laureys S (eds.), *Neuroimaging of Consciousness*. Berlin, Heidelberg: Springer, 2013, 13382.
- Nir Y, Vyazovskiy VV, Cirelli C, et al. Auditory responses and stimulus-specific adaptation in rat auditory cortex are preserved across NREM and REM Sleep. *Cereb Cortex* 2013;23:1362–78.
- Oudiette D, Paller KA. Upgrading the sleeping brain with targeted memory reactivation. *Trends Cogn Sci* 2013;17:142–9.
- Perrin F, Garcia-Larrea L, Mauguier F, et al. A differential brain response to the subject's own name persists during sleep. *Clin Neurophysiol* 1999;110:2153–64.
- Polich J. Updating P300: an integrative theory of P3a and P3b. *Clin Neurophysiol* 2007;118:2128–48.
- Pulvermüller F, Kujala T, Shtyrov Y, et al. Memory traces for words as revealed by the mismatch negativity. *NeuroImage* 2001;14:607–16.
- Rasch B, Born J. About sleep's role in memory. *Physiol Rev* 2013;93:681–766.
- Reed JM, Squire LR. Impaired recognition memory in patients with lesions limited to the hippocampal formation. *Behav Neurosci* 1997;111:667–75.
- Rosenthal CR, Andrews SK, Antoniadis CA, et al. Learning and recognition of a non-conscious sequence of events in human primary visual cortex. *Curr Biol* 2016;26:834–41.
- Rosenthal CR, Kennard C, Soto D. Visuospatial sequence learning without seeing. *PLoS ONE* 2010;5:e11906.
- Rotello CM, Macmillan NA, Reeder JA. Sum-difference theory of remembering and knowing: a two-dimensional signal-detection model. *Psychol Rev* 2004;111:588–616.
- Ruby P, Caclin A, Boulet S, et al. Odd sound processing in the sleeping brain. *J Cogn Neurosci* 2008;20:296–311.

- Rugg MD, Mark RE, Walla P, et al. Dissociation of the neural correlates of implicit and explicit memory. *Nature* 1998;**392**:595–8.
- Scott RB, Dienes Z, Barrett AB, et al. Blind insight: metacognitive discrimination despite chance task performance. *Psychol Sci* 2014;**25**:2199–208.
- Smulders FT, Miller JO, Luck SJ, et al. The lateralized readiness potential. In: *Oxford Handbook of Event-Related Potential Components*. Oxford: Oxford University Press; 2012, 209–29.
- Squire LR, Wixted JT, Clark RE. Recognition memory and the medial temporal lobe: a new perspective. *Nat Rev Neurosci* 2007;**8**:872–83.
- Strauss M, Sitt JD, King JR, et al. Disruption of hierarchical predictive coding during sleep. *Proc Natl Acad Sci* 2015;**112**:E1353–62.
- Tononi G, Cirelli C. Sleep and the price of plasticity: from synaptic and cellular homeostasis to memory consolidation and integration. *Neuron* 2014;**81**:12–34.
- Tulving E. Memory and consciousness. *Can Psychol Can* 1985;**26**:1–12.
- Vilberg KL, Rugg MD. Memory retrieval and the parietal cortex: a review of evidence from a dual-process perspective. *Neuropsychologia* 2008;**46**:1787–99.
- Voss JL, Paller KA. Neural correlates of conceptual implicit memory and their contamination of putative neural correlates of explicit memory. *Learn Mem Cold Spring Harb N* 2007;**14**:259–67.
- Webb WB. Letters to the Editor. *Psychol Sci* 1990;**1**:329–.
- Wood JM. 1990. *Implicit and Explicit Memory for Verbal Stimuli Presented During Sleep*. The University of Arizona.
- Wood JM, Bootzin RR, Kihlstrom JF, et al. Implicit and explicit memory for verbal information presented during sleep. *Psychol Sci* 1992;**3**:236–9.