The development of long-term lexical representations through Hebb repetition learning

Arnaud Szmalec [a,*], Mike P.A. Page [c], Wouter Duyck [b]

[a] Université Catholique de Louvain, Louvain-la-Neuve, Belgium
[b] Ghent University, Ghent, Belgium
[c] University of Hertfordshire, Hatfield, UK

ARTICLE INFO

ABSTRACT

This study clarifies the involvement of short- and long-term memory in novel word-form learning, using the Hebb repetition paradigm. In Experiment 1, participants recalled sequences of visually presented syllables (e.g., *la-va-bu-sa-fa-ra-re-si-di*), with one particular (Hebb) sequence repeated on every third trial. Crucially, these Hebb sequences contained three orthographic nonword neighbors of existing Dutch base-words (e.g., *lavabu – lavabo* [*kitchen sink*]). Twenty-four hours later, the same participants performed two auditory lexicalization tests involving the actual Dutch base-words (e.g., *lavabo*, *safari*, *residu*). Both tests yielded slower reaction times for these Dutch base-words compared with matched control words, which reflects lexical competition between the base-words and the Hebb sequences, therefore demonstrating lexical engagement of the Hebb sequences. In Experiment 2, we subsequently used the Hebb paradigm as an analogue of word-form learning, in order to investigate whether the creation of novel lexical memories requires sleep. Whereas earlier findings indicate that overnight sleep plays a crucial role in lexical consolidation, the current results show that Hebb learning of phonological sequences creates novel word-forms representations in the mental lexicon by the mere passage of time, with sleep playing no necessary role.

© 2012 Elsevier Inc. All rights reserved.

Introduction

Novel word-form learning

It has been hypothesized that verbal short-term memory is primarily a language learning device (e.g., Baddeley, Gathercole, & Papagno, 1998; Duyck, Szmalec, Kemps, & Vandierendonck, 2003; Gathercole, Hitch, Service, & Martin, 1997; Gupta, 2003; Page & Norris, 2008, 2009). In this view, the ability to temporarily retain either verbal information in daily life (like a telephone number) or a sequence of nonsense syllables in a memory experiment, is simply a by-product of its primary evolutionary purpose, that is, the acquisition of language (or, more specifically, of phonological word-forms). The involvement of short-term memory processes in language acquisition is supported by evidence in a large body of correlational findings, showing a strong association between measures of immediate verbal serial recall (e.g., verbal span) and measures of experimental word learning (e.g., nonword paired-associate learning), both in children and adults. Furthermore, a variety of data from neuropsychological case-studies, from people with learning disabilities, and from gifted language learners, supports the position that short-term serial recall and word-form acquisition are functionally related (see Baddeley et al. (1998) for a review). In more recent years, evidence has emerged that in particular short-term memory for serial order information, rather than for item information, is a strong predictor of word-form learning, both in adults and children (e.g., Leclercq & Majerus, 2010; Majerus, Poncelet, Elsen, & Van der Linden, 2006;

* Corresponding author. Address: Psychological Sciences Research Institute, Université Catholique de Louvain, 10 Place du Cardinal Mercier, 1348 Louvain-la-Neuve, Belgium. Fax: +32 (0)9 264 64 96.
E-mail address: arnaud.szmalec@UGent.be (A. Szmalec).

Majerus, Poncelet, Greffe, & Van der Linden, 2006; Majerus, Poncelet, Van der Linden, & Weekes, 2008).

Although the association between immediate serial recall and word-form learning is a well-established finding, it has proved difficult to make explicit how the mechanisms responsible for the short-term retention of verbal serial information exactly map onto those responsible for the learning of novel phonological word-forms. Quite independently from these developments in the working memory literature, researchers have been dealing with similar theoretical issues within the domain of sequential learning, trying to clarify the relationship between sequential learning and language learning for several years (e.g., Conway & Christiansen, 2001; Conway & Pisoni, 2008). At present, however, the exact involvement of working memory in language/word learning remains underspecified. In order to resolve this theoretical gap, short-term memory theorists have recently tried to incorporate data on word-form learning into existing computational models of immediate serial recall (e.g., Burgess & Hitch, 2006; Gupta, 2008; Page & Norris, 2008, 2009). Those computational modeling efforts have led to the hypothesis that the same memory mechanisms, and more precisely those specifically responsible for representing serial order information, underlie both immediate serial recall and the acquisition of novel word-forms. It is reasoned that a newly acquired word-form is basically a familiarized sequence of discrete elements (like letters, phonemes or syllables). This implies that learning a sequence of letters, like B J F M L, in a short-term memory experiment, is functionally equivalent to learning the word-form "bejayeffemmelle" (Page & Norris, 2009), that can then be mapped onto its meaning. Although these hypotheses are seemingly straightforward and plausible, and although performance in recall tasks fits these computational models well, there is hitherto very little experimental evidence demonstrating that naturalistic learning of a novel phonological word-form essentially comes down to the learning of a grouped sequence of phonological items. The present study aims to provide this empirical evidence, necessary to validate and further refine the Page and Norris (2008, 2009) unifying modeling framework of immediate serial recall and word-form learning.

Novel word-form learning and the Hebb repetition effect

According to Page and Norris (2008, 2009), the Hebb repetition paradigm stands at the interface between immediate serial recall and language learning. Hebb (1961) asked participants to perform an immediate verbal serial recall task in which one particular sequence of items was repeated every third trial. He observed that recall for repeating sequences increased substantially compared with non-repeating sequences, a phenomenon which is known as the Hebb repetition effect. In essence, the Hebb effect is a serial order-learning effect which shows how a retained sequence of information in short-term memory gradually develops into a stable long-term memory trace. In their connectionist modeling framework, which is based on their primacy model of short-term memory (Page & Norris, 1998), Page and Norris (2008, 2009) proposed that the Hebb repetition effect, which is a paradigmatic

example of long-term sequence learning, can be seen as a laboratory analogue of naturalistic word-form acquisition, in as much as novel word-forms are themselves grouped sequences of familiar sublexical items. The hypothesis that the Hebb effect mimics naturalistic word-learning has so far been tested in only two recent studies. The first one is a correlational study by Mosse and Jarrold (2008), in which the authors demonstrated Hebb repetition learning in 5- and 6-year old children, both in verbal and spatial immediate serial recall tasks. Crucially, they observed that the magnitude of Hebb learning, in both modalities, correlated significantly with nonword paired-associate learning (used as a proxy of word learning), but not with word paired-associate learning. This study was the first to show an association between Hebb learning and novel word-form learning, suggesting that the degree of sequence learning observed in the Hebb effect is related to the degree of sequence learning when acquiring novel word-forms. This conclusion was further elaborated in an experimental study by Szmalec, Duyck, Vandierendonck, Barberá Mata, and Page (2009). Participants saw sequences of nonsense syllables, presented one at a time. The nine-item sequences (e.g., zi-lo-ka-ho-fi-se-be-ru-mo) were grouped by pauses into three groups of three syllables and certain groups were repeated throughout the experiment consistent with, though not identical to, Hebb's (1961) learning procedure. In a subsequent experiment, that followed shortly after the first, participants performed an auditory lexical decision task on nonwords that had been constructed from the ordered syllables taken from the previously repeated Hebb sequences (ziloka, hofise, berumo). Interestingly, these Hebb-based nonwords, that had been presented visually, yielded slower auditory lexical decision times than control nonwords, that is, they were slower to be identified as nonwords. This demonstrated that the repeated sequences of syllables, which were learned in a visual Hebb-like procedure, established novel phonological word-forms in lexical memory.

Gaskell and Dumay (2003) have argued that many tests of whether a new word has been learnt do not necessarily measure the lexicalization process in itself. They can, for instance, indicate whether a particular word is familiar, but they do not unambiguously address whether or not the information has been stored in the mental lexicon. This differentiation between mere phonological learning vs. lexical integration was further empirically and theoretically elaborated by Leach and Samuel (2007), who distinguish two processes in the concept of lexicalization, namely lexical configuration and lexical engagement. Lexical configuration refers to acquiring the factual knowledge that is associated with a word (e.g., its sound, meaning or spelling), whereas lexical engagement refers to the interaction of a novel word-form with existing entries in the mental lexicon. In the current study we adopted this more stringent aspect of lexicalization, investigating whether Hebb learning of verbal sequences not only leads to lexical configuration but also to lexical engagement, and therefore mimics naturalistic vocabulary acquisition, as proposed in the Page and Norris (2008, 2009). This approach is derived from Gaskell and Dumay (2003), according to whom lexical engagement of an item can be assessed by measuring

its ability to affect the activation of other representations that are already stored in the mental lexicon, a phenomenon known as lexical competition (see also Dumay & Gaskell, 2007; Fernandes, Kolinsky, & Ventura, 2009; Leach & Samuel, 2007). In their study, Gaskell and Dumay (2003; Experiment 3) familiarized participants with pseudowords (e.g., cathedruke) that strongly overlapped with existing words (e.g., cathedral). They observed that familiarization of these pseudowords (i.e. the created lexical competitors) did not alter the recognition of the existing words (i.e. the base-words) in a pause detection task that was administered immediately after the familiarization. However, 1 week later and without any further exposure to the pseudowords, a clear lexical competition effect emerged, which suggests that lexical engagement requires a consolidation period (Bowers, Davis, & Hanley, 2005; Dumay & Gaskell, 2007). The inhibitory effect of lexical competition on pause detection was explained by Gaskell and Dumay as follows: In a pause detection task, participants are instructed to detect a short pause that is artificially embedded in connected speech. Mattys and Clark (2002) demonstrated that the speed with which this artificial pause can be detected, depends on the overall amount of lexical activity caused by the preceding speech (cohort activation), making it a good test of lexical activation/access. For example, words with a late uniqueness point (e.g., blackberry) that have a pause inserted near the end of the word (blackb_erry), will activate several lexical representations (e.g., blackbox, blackbird, blackboard, etc.) during processing of the onset syllables. This activation of multiple lexical candidates is assumed to consume processing resources that could otherwise be allocated to the detection of the pause. Therefore, the speed at which the pause can be detected is a function of the number of phonological neighbors (or, by extension, lexical competitors) of the target word, stored in the mental lexicon. The lexical competition paradigm was initially constructed using lexical competitors that greatly overlap with existing words (e.g., alcohol vs. alcohin in Gaskell and Dumay (2003)). More recently, Dumay and Gaskell (2012) observed equally strong lexical competition between words and nonwords that are less closely related (e.g., muck vs. lirmucktoze). By showing that the lexical competition effect is not restricted to close variants of existing lexical entries, the Dumay and Gaskell (2012) study endorses the lexical nature of the effect.

Another nice demonstration of the usefulness of lexical competition as a measure of lexical engagement, is the study by Fernandes et al. (2009). They used the well-established artificial language learning paradigm (ALL; Saffran, Aslin, & Newport, 1996; Saffran, Newport, & Aslin, 1996) and found that novel word-forms that are extracted from a continuous artificial language stream based on statistical regularities, cause interference with existing words within the mental lexicon, as measured through lexical decision. By doing so, they are the first to directly demonstrate that the output of statistical segmentation processes during artificial language learning shows lexical engagement in a manner also consistent with the framework proposed by Page and Norris (2008, 2009).

The role of sleep in novel word-from learning

The present study proposes the Hebb repetition paradigm as an artificial analogue of lexical learning, with the aim of understanding the memory processes that drive the acquisition of novel words. One central question related to word-form acquisition that is receiving increasing attention over recent years, is whether or not sleep plays a crucial role in long-term lexical learning (e.g., Davis, Di Betta, Macdonald, & Gaskell, 2009; Dumay & Gaskell, 2005, 2007; Tamminen, Payne, Stickgold, Wamsley, & Gaskell, 2010). Dumay and Gaskell (2007; see also Davis et al., 2009) asked participants to monitor a sequence of pseudowords for the occurrence of a specific phoneme (i.e., phoneme monitoring), one half of the participants in the morning, the other half in the evening. They subsequently tested whether the exposure to these pseudowords resulted in novel word-form representations by testing lexical competition 12 h later. That is, after a day being awake for the morning group and after a regular night of sleep for the evening group. Interestingly, they found support for lexicalization of the novel word-forms in the evening group but not in the morning group. The morning group did however eventually also show lexicalization effects when they were tested again 24 h after the familiarization phase, when they had slept. Overall, these findings support the idea that sleep is crucial for the lexicalization of new verbal memories. Tamminen and his colleagues (2010) more recently examined the neurophysiological sleep processes that drive this offline consolidation of lexical memories, using a procedure similar to that of Dumay and Gaskell (2007). In a nicely designed experiment, they showed that sleep spindle activity (11–15 Hz oscillations in the sleep EEG) predicts the strength of overnight lexical integration of novel word-forms, given that there was a .59 correlation between number of spindles and the magnitude of the lexical competition effects. These findings are consistent with the so-called two-stage memory account of novel word learning which proposes that there are two complementary learning systems, one at the level of the hippocampus and one at the level of the neocortex (Davis et al., 2009; McClelland, McNaughton, & O'Reilly, 1995; O'Reilly & Norman, 2002). Within this view, new verbal memories are initially held in hippocampal networks, from where they are gradually transferred to long-term neocortical networks, and it is the neocortical learning in particular that is assumed to rely on sleep.

Although the available empirical evidence favors the view that the consolidation of newly acquired word-forms requires sleep, the generalizability of this finding across different word learning paradigms remains an interesting debate. This concern is based on what is currently known about the interaction between sleep and human memory in general, of which lexical memory is only one instance. Early empirical evidence for the role of sleep in memory consolidation is almost a hundred years old (Jenkins & Dallenbach, 1924). Decades of research have subsequently shown that sleep enhances consolidation in procedural and declarative memories, but whether or not sleep shows its beneficial effects very much depends on the specific context in which the novel memories have been acquired

(see Diekelmann, Wilhelm, and Born (2009) for a review). Of particular interest to the current study, is Diekelmann et al.'s (2009) conclusion that the sleep advantage is greater for shallow and explicitly learned memory traces. If interaction effects between learning characteristics and sleep in long-term memory consolidation exist in general, one would also expect the same to be true for the consolidation of lexical memories. Consolidation of novel word-forms may benefit from sleep, as demonstrated in the Dumay and Gaskell (2007) approach, but it is unclear whether sleep is a critical factor: is it a *conditio sine qua non* for lexical consolidation, or do the favorable effects of sleep depend on the learning conditions? We will investigate this issue by applying the sleep consolidation rationale to our word learning paradigm derived from the Hebb repetition effect.

The present study reports two experiments. Experiment 1 seeks to make the interaction between short-term memory (for serial-order information) and language learning more explicit by focussing on the question whether the Hebb repetition effect for verbal materials mimics phonological word-form learning. Therefore, participants recalled sequences of syllables following a Hebb learning protocol. Twenty-four hours later, the same participants were involved in two lexicalization tests, in order to investigate whether learning the Hebb sequences had resulted in the creation of a novel lexical entry, just like novel word-forms do. The main goal of Experiment 2 is to investigate whether earlier conclusions about the crucial involvement of sleep in consolidating novel word-forms (e.g., Dumay & Gaskell, 2007; Tamminen et al., 2010) may be generalized to our newly introduced word acquisition paradigm, based on Hebb repetition learning.

Experiment 1

In order to test the hypothesis that verbal Hebb repetition learning can be used as an analogue of naturalistic word-form learning, participants in Experiment 1 first learned ordered sequences of nine nonsense syllables following a (nearly) standard Hebb learning protocol, that is, with the contents of one particular sequence (e.g., *la-va-bu-sa-fa-ra-re-si-di*) repeated every third trial. Crucially, these Hebb sequences contained three (three-syllable) nonwords that constitute lexical neighbors of existing Dutch base-words. Twenty-four hours later, the same participants were required to perform a pause detection task and an auditory lexical decision task involving the Dutch base-words for which a lexical competitor was hypothesized to be created during Hebb learning in the first experiment (e.g., *lavabo, safari, residu*). Pause detection on the base-words is assumed to be a stringent test of lexicalization because it measures lexical engagement (Leach & Samuel, 2007), i.e. the interaction of the mental representation that the Hebb sequence is purported to create with previously consolidated (known) target words, without requiring any lexicality judgement on the new word itself. For the sake of completeness, we still also included lexical decision in this experiment in order to see whether both measures converge in the present Hebb-learning context. Both lexicalization tests took place 24 h later, given that Dumay and Gaskell (2007) showed that engagement in

lexical competition requires a consolidation period that should involve sleep (see also Davis et al., 2009; Dumay & Gaskell, 2005). Our predictions were as follows: if Hebb-repetition-based learning of the phonological materials generates new entries in the mental lexicon, just like naturalistic word-learning does, we anticipate lexical competition to have arisen within 24 h, between the implicitly learned syllable sequences and the Dutch base-words. Consequently, we predict slower pause detection and longer lexical decision times for those Dutch base-words, compared with a set of control words that do not have such implicitly learned lexical competitors.

Method

Participants

A total of 58 adults (36 females), aged between 18 and 26 years (mean 20.6 years) were paid €20 for participation. They were all Dutch speaking and naive to the purpose of the study.

Materials

Hebb learning. Sequences of nine syllables, all consonant-vowel structures (CVs), were presented to the participants for immediate serial recall. Each participant completed two sessions of 36 sequences. Within one session, the contents of a given (Hebb) sequence (though not the exact sequence – see below) were repeated every third trial, which implies that we had 24 unrepeated (or filler) sequences and one Hebb sequence whose content was repeated 12 times. The Hebb sequences consisted of three three-syllable groupings that overlapped with existing Dutch words. The sequence *la-va-bu-sa-fa-ra-re-si-di,* for example, consists of three three-syllable subgroups, namely *la-va-ba, sa-fa-ra* and *re-si-di*, which overlap with the Dutch base-words *lavabo, safari* and *residu*. The entire list of words that were used is represented in Table 1. The order of the CVs within the three-syllable subgroups

Table 1
CVCVCV syllable sequences and overlapping base-words with phonemic transcription and English translation.

CVCVCV sequence	Base-word	Transcription	English Translation
bi-ki-na	bikini	/bikini/	*Bikini*
fi-na-lo	finale	/finalə/	*final*
fy-si-cu	fysica	/'fizika/	*physics*
ho-re-co	horeca	/horeka/	*catering*
ka-ra-to	karate	/ka'ratə/	*karate*
la-va-bu	lavabo	/lava'bo/	*kitchen sink*
la-wi-na	lawine	/lawinə/	*avalanche*
li-bi-du	libido	/'libido/	*libido*
me-ri-tu	merite	/me'ritə/	*merit*
no-ma-di	nomade	/no'madə/	*nomad*
pa-ra-di	parade	/paradə/	*parade*
re-si-di	residu	/rezi'dy/	*residue*
sa-fa-ra	safari	/safari/	*safari*
sa-la-du	salade	/saladə/	*salad*
sa-la-mo	salami	/sɑ'lami/	*salami*
sa-ti-ra	satire	/sɑtirə/	*satire*
va-li-do	valide	/va'lidə/	*valid*
vi-si-ti	visite	/vi'zitə/	*visit*

was kept constant, but not the order of the entire Hebb sequence. For example, a legal Hebb "repetition" of the sequence above could be *re-si-di-la-va-bu-sa-fa-ra*. This is a more conservative approach than a typical Hebb procedure, in which all nine CVs would be presented in exactly the same order at each repetition trial. Because using partial repetitions is likely to counteract Hebb learning, this adds strength to any Hebb effect yielded by this procedure. The Hebb repetition effect is meant to be a pure measure of serial-order learning (i.e., not confounded with item learning). Therefore, the unrepeated, filler sequences were constructed from the same CVs as the Hebb sequences, the only difference being that the order of the CVs was randomly determined on filler trials, whereas it was fixed (in three three-syllable subgroups) on Hebb trials.

Lexicalization tests. The critical materials for the lexicalization tests were 18 trisyllabic Dutch base-words (of the form CVCVCV) that overlapped with the 18 nonwords that constituted the repeating Hebb sequences. The Dutch base-words differed from these Hebb sequences in their final vowels (see Table 1), so that the words' uniqueness points were as late as possible (relative to the Hebb sequence); this is known to increase the sensitivity to lexical competition effects (Gaskell & Dumay, 2003). Only words that had no existing lexical neighbors in Dutch were chosen, in order to maximize potential interference effects of the newly learned lexical competitor (i.e., the Hebb sequence). The 18 base-words had a mean frequency of 2.77 (occurrences per million, as per Duyck, Desmet, Verbeke, & Brysbaert, 2004). Because two Hebb lists were learnt in this experiment and because three 3-syllable words can be constructed from one Hebb-list, we could only use 6 (out of the 18) base-words per participant. Therefore, for each participant, 6 base-words were selected from the set of 18 in a counterbalanced fashion so that each base-word was used a comparable number of times over the entire experiment. Following the rationale of Hebb learning of lexical competitors, the 6 base-words (e.g., *lavabo*) thus overlapped with the syllable sequences that were learned in the Hebb task (*la-va-bu-*…). Of the remaining 12 base-words, 6 were matched (on frequency and neighbors) and used a control words for the lexicalization tests. Finally, the remaining 6 base-words were used as fillers in the lexicalization tests. This counterbalancing procedure also implies that the same Dutch words constitute the control condition for some participants, but the lexical competition condition for an equal number of other participants, thereby ensuring stimulus matching across conditions.

In the *pause detection* (PD) task, 100 words were used. Twenty-five of these had a CVCVCV structure: the base-words and control words described above, and filler words. There were also 25 fillers with a different structure (e.g., *kalmte:* calmness). The 50 words altogether were presented once without (pause-absent trials) and once with an artificially embedded 150 ms pause (pause-present trials), in a random fashion. The silent pauses were placed in the WAV files using the same sound editing software. For the bisyllabic words, the pause was inserted before the final syllable (e.g., *kalm_te*). For the trisyllabic words, the pause was inserted just after the second syllable (e.g., *ali_-nea*) or just before the final syllable (e.g., *domi_no*). As a

result, the words without pause were 800 ms long, and those with a pause were 950 ms long. Words for lexical decision and pause detection were presented auditorily, in contrast with the stimuli for the Hebb experiment, which were presented visually. This excludes episodic memory influences and supports the somewhat abstract phonological nature of the representations assumed to be involved in the lexical competition.

In the *lexical decision* (LD) task, a total of 100 words were presented, 50 words and 50 nonwords. The words included the 6 critical base-words for that given participant, the 6 control words, 13 CVCVCV filler words that were not matched to the base-words (including the remaining words from the critical set of 18), and 25 filler words that were not of the form CVCVCV, like *dakpan* (tile). The nonwords included 25 CVCVCV items (e.g., *kaluwo*) and 25 nonwords that were not of the CVCVCV format (e.g., *schrak*). All stimuli for the LD task were digitally recorded in WAV format, pronounced by a female speaker. In order to match presentation times, we edited the WAV files in sound editing software (WaveLab) and transformed them into files of exactly 800 ms, without any audible loss of quality.

It is important to realize that the critical stimuli for the auditory lexicalization tests were the base-words (e.g., *safari*) and that the nonsense Hebb sequences (*e.g., sa-fa-ra*) were never presented as word-forms to the participants. One potential complication that is inherent to the lexical competition method is that we cannot control the pronunciation of the phonological forms that are acquired through Hebb learning of visual syllable sequences. However, all Dutch syllables involved in this experiment are pronounced the same in isolation as included in the associated word-form. In addition, if the phonological representation of some nonwords would not match the standard pronunciation, this would minimize the overlap between the base-words and the lexical competitors and hence work against our hypothesis.

Procedure

The Hebb learning procedure was as similar as possible to that of Page, Cumming, Norris, Hitch, and McNeil (2006). The CVs were presented serially and visually. They remained on the screen for 500 ms. Immediately after presentation, a recall screen was presented in which the nine CVs were arranged randomly in a "noisy" circle around a central question mark. Participants were instructed to recall the CVs in the same order as they were presented by clicking with a computer mouse on each item in turn. The question mark was clicked to indicate an omission, at the position in the sequence were the omission occurred. This way, potential correct responses after an omission are still in the right serial position and can thus be counted as correct. The Hebb learning experiment lasted approximately 75 min.

All participants were clearly instructed about the requirement to go through at least 6 h of sleep before taking part in the second part of the experiment where the lexicalization of the Hebb sequences was investigated. Participants that appeared to have slept less than 6 h were excluded. They were not informed about any possible

relation between the Hebb learning and lexicalization stages of the experiment. The stimuli for the PD and LD tasks were presented through closed headphones (Sennheiser HD 265-1) at 60 dB. In the PD task, the presentation time of the stimuli was 800 or 950 ms, depending on the inclusion of a pause, also followed by a fixed 2500 ms interstimulus interval. Participants were instructed to respond as fast and accurately as possible whether an artificial pause was embedded in the word or not, by pressing a key on the response box. Reaction times in the PD task were measured from the onset of the embedded pause. In the condition without a pause, reaction times were measured from the same time point in the digitized speech file as in the pause condition. In the LD task, the presentation time of the stimuli was 800 ms, followed by a fixed interstimulus interval of 2500 ms. Participants were required to decide as fast and accurately as possible whether the stimulus was a word or a nonword, by pressing a key on the response box. The order of both lexicalization tests was counterbalanced across participants. The procedure for the lexicalization tests lasted approximately 20 min.

Results

Hebb learning

According to the standard Hebb learning protocol, a CV was scored as correct if it was recalled in the correct position in the sequence. Mean proportions of correctly recalled CVs by trial number for the Hebb and the filler sequences are displayed in Fig. 1. The Hebb repetition effect was measured by taking the gradient of the regression line linking the performance on successive "repetitions" of the Hebb sequences and comparing it with the corresponding gradient for the filler sequences, for each individual participant. The gradient values were entered into an analysis of variance with sequence type (filler vs. Hebb) as the independent variable for testing the Hebb repetition effect. The results show that the gradient for the filler sequences ($M = .005$, $SE = .008$) was significantly lower than that for the Hebb sequences ($M = .020$, $SE = .016$): $F(1,57) = 53.61$, $n_p^2 = .48$, $p < .001$. These results show that a clear Hebb effect was obtained, which is a necessary condition for considering the results of the lexicalization tests.

Lexicalization tests

All participants met the 6 h sleep requirement; the average amount of nocturnal sleep was 7 h and 18 min. Mean RTs and accuracy in the different stimulus conditions of the PD task and the LD task are presented in Table 2. Because only the difference between the base-words, i.e. those that overlap with the Hebb sequences, and the matched control words are of theoretical interest, we only report the planned comparisons assessing the lexical competition effect (e.g., Dumay & Gaskell, 2007). Regarding the pause detection results,[1] we did not observe any reliable difference between pause-present and pause-absent trials, so the data were averaged across both trial types (see Dumay & Gaskell, 2007). The pause detection times were slower for the base-words than for the control words, $F_1(1,57) = 7.24$, $n_p^2 = .11$, $p < .01$; $F_2(1,17) = 5.99$, $n_p^2 = .26$, $p < .05$. The accuracy data for the pause detection task did not reveal significant differences between the base and control words, F_1 and $F_2 < 1$. The analyses further reveal that the lexical decision times for the base-words were reliably slower than for the control words, $F_1(1,57) = 6.23$, $n_p^2 = .10$, $p < .05$; $F_2 < 1$. The accuracy of the lexical decisions was comparable for base-words and control words, F_1 and $F_2 < 1$.

Discussion

The aim of Experiment 1 was to investigate the lexical status of phonological materials that are acquired through Hebb learning. Participants recalled visually presented sequences of nine syllables, following a Hebb repetition learning procedure in which three three-syllable subgroups were repeated across lists. Exactly 24 h later, the same participants performed an auditory lexical decision task and a pause detection task on a subset of Dutch words that have no orthographic neighbors except the three-syllable nonwords that were learned in the preceding Hebb learning experiment. Interestingly, the results show that participants were reliably slower to make a lexical decision on those Dutch base-words, compared with a subset of matched control words for which no lexical competitor was learned in the Hebb experiment. Moreover, in the pause detection task (Mattys & Clark, 2002), participants were slower to detect whether or not a 150 ms pause was artificially embedded in the same Dutch base-words. Thus, in line with our predictions, the results of both lexicalization tests show lexical engagement, 24 h after Hebb sequence learning, reflecting a competition between the learned syllable sequences and the Dutch base-words. Note that the same base-words served in the lexical competition (Hebb) condition for one half of the participants, but as control words for the other half of participants. So, across participants, RTs reflect pause detection and lexical
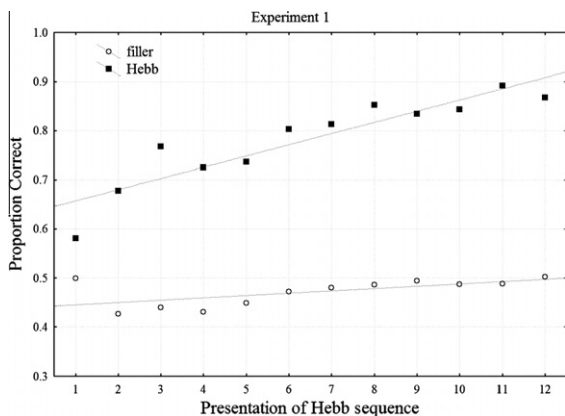

Fig. 1. Accuracy (proportion correct) for Hebb and filler sequences in Experiment 1. Regression lines have been added to show the improvement in performance. Values for filler trials represent the average of the two filler sequences that were presented in between each Hebb repetition.

---

[1] With respect to the analyses by items (F_2), we note that our 18 items are not a random sample of the population but are all the available CVCVCV words in the Dutch language that comply with the stimulus requirement for this study.

Table 2
Mean response times and accuracies in the different conditions of the pause detection task and the lexical decision task used in Experiment 1. Standard deviations are in parentheses.

	RT (ms)	Accuracy (%)
Pause detection		
Base-words	720 (*150*)	90 (*11*)
Control words	692 (*129*)	89 (*11*)
Filler words	557 (*118*)	85 (*10*)
Lexical decision		
Base-words	853 (*123*)	90 (*15*)
Control words	828 (*105*)	91 (*16*)
Filler words	832 (*98*)	91 (*13*)
Nonwords	850 (*103*)	90 (*19*)

decision latencies to exactly the same word stimuli. Overall, the findings of Experiment 1 indicate that phonological word-form learning relies on the cognitive processes responsible for representing serial-order information in memory, of which Hebb repetition learning is a paradigmatic example. Therefore, it can be concluded that Hebb repetition learning is a laboratory analogue of naturalistic word-form acquisition, which is the integration of newly acquired word memories in the mental lexicon.

Experiment 2

The finding that novel word-forms acquired through Hebb repetition learning show lexical competition effects after a regular period of nocturnal sleep is in line with earlier studies that showed that the lexical consolidation of novel words requires sleep (e.g., Dumay & Gaskell, 2007; Tamminen et al., 2010). Because sleep was not manipulated in Experiment 1, it is not possible however to draw firm conclusions about the relation between sleep activity and the lexicalization of verbal Hebb sequences. The relation between sleep and lexical consolidation after Hebb learning is however particularly relevant to address because it has been shown in the literature (see Diekelmann et al. (2009) for a review) that the role of sleep in the consolidation of memories is dependent on how these memories have been acquired. In this view, the question whether consolidation requires sleep is not a function of *what* is learned (here: words), but *how*. Clearly, there are differences between our approach and the one by Dumay and Gaskell (2007) in this respect. Whereas we used the Hebb paradigm, participants in the Dumay and Gaskell (2007; and also in Davis et al. (2009) and Tamminen et al. (2010)) were exposed to the novel word-forms through a phoneme-monitoring task. During our Hebb repetition learning, we hypothesize that participants implicitly create novel lexical memories by picking up regularities in the phoneme sequences. In the Gaskell and Dumay paradigm, participants are explicitly made aware of the fact that they have to try to learn novel words and that their memory for those words will be tested afterwards. In order to make sure that the novel words are indeed processed, they are required to listen to the words and detect the presence of a pre-specified phoneme.

Diekelmann et al. (2009) identify several learning characteristics that are relevant for the issue of sleep-based memory consolidation, of which two deserve to be further elaborated in the light of the present study. The first difference relates to the *strength of the memory trace*. Diekelmann et al. (2009) review a number of studies which indicate that "benefits from sleep are greater for weaker than stronger traces" (p. 313). The phoneme monitoring task is known to be a rather shallow nonword exposure task (Gaskell & Dumay, 2003). It does not involve recall of the memoranda like the Hebb procedure does, so it is likely that the memory traces of the phoneme sequences created during phoneme monitoring are not as strong as the representations of the word-forms acquired under Hebb repetition learning.

A second relevant determinant of sleep involvement in memory consolidation is the *implicitness* of the acquired memories. According to Diekelmann et al.'s (2009) literature review, "sleep preferentially consolidates explicitly learned materials" (p. 314). In the phoneme-monitoring task, instructions explicitly state that nonsense words will be presented for phoneme monitoring and that memory for these words will be measured later. By contrast, no word-forms are presented in a Hebb-learning protocol; participants are instructed to attend to a sequence of visually presented syllables for immediate serial recall, without any instruction about, nor reference to, word-form learning at all. Because participants are made aware of the fact that they are learning novel words during phoneme monitoring, but not during Hebb learning, the novel lexical memories are acquired in a more implicit way in the Hebb repetition paradigm, similar to the way children implicitly acquire novel lexical representations through echoing regularities in the phonological input from their environment.

There are clearly more differences between phoneme monitoring and Hebb sequence learning than we have discussed here but, as we argued above, the nature of some of these differences makes it worthwhile to question whether the presumed crucial role of sleep in lexical consolidation may also be supported in the context of Hebb repetition learning. To address this question, an experiment was designed in which participants learned sequences of syllables in the same way as in Experiment 1, before examining lexical engagement of the Hebb sequences on three different occasions within a 24-h time span. For half of the participants (hereafter called the *morning group*), Hebb learning took place in the morning (between 8 and 10 a.m.) and the lexicalization test was administered a first time immediately after Hebb learning, a second time 12 h later (i.e. after 12 h wakefulness) and, finally, a third time 24 h after Hebb learning (i.e. after a normal period of nocturnal sleep). For the second half of the participants (the *evening group*), Hebb learning took place in the evening (between 8 and 10 p.m.) and the lexicalization test was administered a first time immediately after Hebb learning, a second time 12 h later (i.e. after a normal period of nocturnal sleep) and finally, a third time 24 h after Hebb learning. This procedure, which is schematically depicted in Fig. 2, allows us to estimate the relative contribution of sleep to the lexicalization of the materials acquired during Hebb learning. If nocturnal sleep (or a sleep-associated factor such as the absence of language input) is crucial to the lexicalization of Hebb sequences, we predict that the morning group will show lexical competition effects only 24 h after Hebb
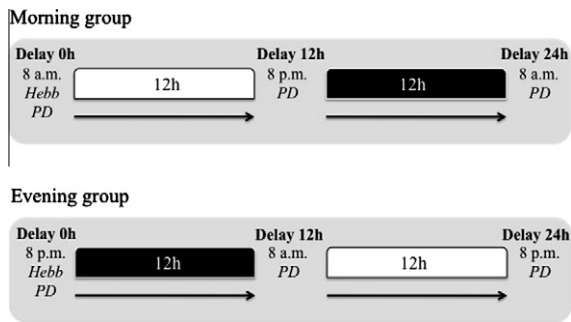
Fig. 2. Schematic representation of Experiment 2. Hebb = Hebb learning task; PD = pause detection task.

learning whereas the evening group would show such effects already 12 h after Hebb learning. If by contrast, the mere passage of time suffices to assimilate Hebb-based word-forms in the mental lexicon, we anticipate that also the morning group will show lexical competition effects 12 h after Hebb learning in contrast to Dumay and Gaskell (2007).

Method

Participants

A total of 92 adults (61 females), aged between 18 and 25 years (mean 20.2 years) were paid €20 for participation. They were all Dutch speaking and naive to the purpose of the study. None of them had participated in Experiment 1.

Materials and procedure

Participants were informed that they would take part in a memory experiment with different tasks spread over 24 h. They were requested to register all their activities within these 24 h in a diary-like report, including their bed and wake-up times. We stressed that a minimum of 6 h of sleep was imperative in the context of this experiment.

The experiment was divided into three parts, based on the time that elapsed between the Hebb learning and the testing of the lexical competition effect, i.e. after a delay of 0 h, 12 h, and 24 h. In the first part of the experiment (8 a.m. and 8 p.m. for the morning and evening groups, respectively), participants went through the Hebb-learning procedure, followed by the pause-detection task (see Fig. 2). This part lasted approximately 80 min. The materials and procedure of both tasks were identical to Experiment 1. Because the lexical-decision and pause-detection tasks yielded comparable results in Experiment 1, we decided to test lexical engagement in Experiment 2 by using the pause-detection task only, similar to Dumay and Gaskell (2007). The second part of the experiment took place 12 h after Hebb learning, that is, after a period of wakefulness for the morning group and regular overnight sleep for the evening group. During this second part, which lasted approximately 15 min, they went through the pause-detection task again. Finally, again 12 h later (24 h after Hebb learning) the pause-detection task was administered once more to both groups (approximately

15 min). Immediately after the final part of the experiment, all participants were subjected to an interview with the experimenter, in which their diaries of the past 24 h were discussed. This interview was especially targeted at examining whether the sleep requirements had been fulfilled and whether the periods of wakefulness had not been contaminated with naps.

Results

Two female participants from the morning group were discarded from the analyses based on the post-experimental interviews, as they reported having taken a nap within the 12 h of required wakefulness prior to their evening tests. The remaining 90 participants (45 per morning/evening group) complied with the experimental protocol. Both groups reported a comparable amount of nocturnal sleep: 7 h and 27 min for the morning group and 7 h and 7 min for the evening group, $t(88) < 1$. For a majority of the participants in both groups, the daily activities had been filled with university classes and study.

Hebb learning

Mean proportions of correctly recalled CVs for the Hebb and the filler sequences, in both the morning and the evening group, are displayed in Fig. 3. The gradient values for the Hebb and filler sequences were entered into an analysis of variance with sequence type (filler vs. Hebb) and group (morning vs. evening group) as the independent variable for testing the Hebb repetition effect in both experimental groups. The analysis reveals a significant main effect of sequence type, $F(1,88) = 99.55$, $n_p^2 = .53$, $p < .001$, whereas the main effect of group was not significant $F(1,88) = 2.35$, $n_p^2 = .03$, $p > .10$. Crucially, the lack of interaction between both factors, $F < 1$, indicates that the Hebb repetition effect was comparable across the morning and evening groups. Further planned comparisons show that in the morning group, the gradient for the filler sequences
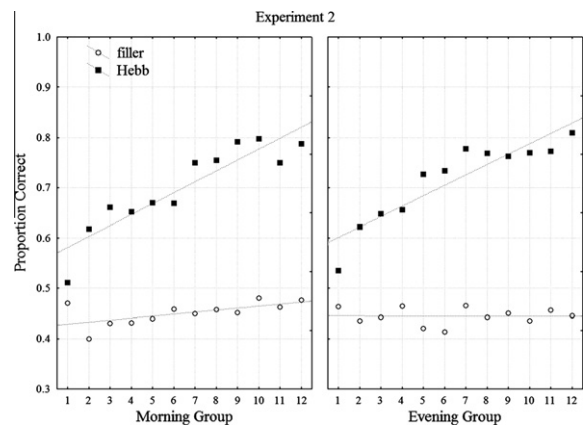

Fig. 3. Accuracy (proportion correct) for Hebb and filler sequences in Experiment 2 for both the morning group (left panel) and the evening group (right panel). Regression lines have been added to show the improvement in performance. Values for filler trials represent the average of the two filler sequences that were presented in between each Hebb repetition.

(M = .004, SE = .009) was significantly lower than that for the Hebb sequences (M = .020, SE = .016), $F(1,88)$ = 42.38, n_p^2 = .32, $p < .001$. In the evening group, the gradient for the filler sequences (M = .000, SE = .008) was significantly lower than that for the Hebb sequences (M = .021, SE = .013), $F(1,88)$ = 57.76, n_p^2 = .40, $p < .001$. Thus, for both groups, a clear and equally large Hebb-learning effect was obtained.

Pause detection

Table 3 presents the mean pause detection times and accuracy, as a function of group (morning group vs. evening group), stimulus condition (base-words vs. control words vs. filler words) and delay (0 h vs. 12 h vs. 24 h after Hebb learning). In order to assess the lexical competition effect, we subtracted the pause detection times for the base-words from the pause detection times for the control words, for each participant individually. Because we did not observe any reliable difference between pause-present and pause-absent trials, the data were again averaged across both trial types (Dumay & Gaskell, 2007). The result of the subtraction, i.e. the lexical competition effect, was entered as the dependent variable in a 2 × 3 ANOVA with group as between subject factor and with repeated measures on the factor delay. The mean values for this ANOVA are graphically represented in Fig. 4. The analyses revealed a significant main effect of delay, $F_1(2,87)$ = 4.57, n_p^2 = .10, $p < .05$; $F_2(2,33)$ = 4.01, n_p^2 = .19, $p < .05$, whereas the main effect of group and the interaction between both factors were not significant, all F's < 1. Further planned comparisons show that there was no lexical competition effect immediately after Hebb learning (delay 0 h), neither for the morning group, nor for the evening group (F's < 1). After a delay of 12 h, both the morning group, $F_1(1,88)$ = 4.77, n_p^2 = .05, $p < .05$; $F_2(1,34)$ = 2.76, n_p^2 = .07, $p = .11$, and the evening group, $F_1(1,88)$ = 5.73, n_p^2 = .06, $p < .05$; $F_2(1,34)$ = 3.55, n_p^2 = .09, $p = .07$, showed a reliable lexical competition effect of 27 ms and 30 ms, respectively. The difference between both groups was not reliable, all F's < 1. Also 24 h after Hebb learning, both the morning group, $F_1(1,88)$ = 9.06, n_p^2 = .09, $p < .01$; $F_2(1,34)$ = 2.69, n_p^2 = .07, $p = .11$, and the evening group, $F_1(1,88)$ = 13.17, n_p^2 = .13, $p < .001$; $F_2(1,34)$ = 4.04, n_p^2 = .11, $p = .05$, showed reliable lexical competition effects of 27 ms and 33 ms, respectively. Again, these competition effects did not differ between both groups, all F's < 1.
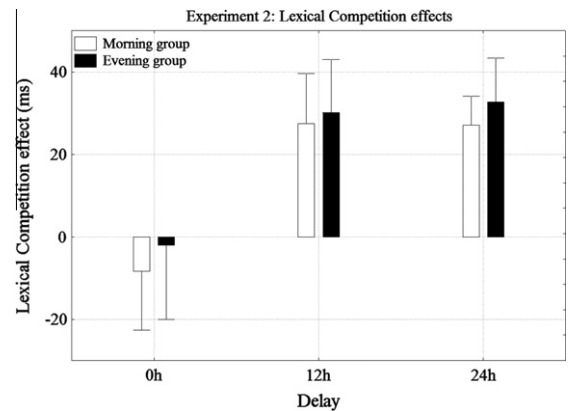

Fig. 4. Lexical competition effect (i.e. pause detection times base-words minus control words) as a function of group (morning group vs. evening group) and delay after Hebb learning (0 h vs. 12 h vs. 24 h) in Experiment 2. Error bars denote standard errors.

The accuracy data from the pause detection task were analyzed using the same 2 × 3 ANOVA with the factors group and delay. The statistical tests did not reveal any significant differences or interactions between groups.

Discussion

The purpose of Experiment 2 was to investigate the role of nocturnal sleep in the lexicalization of novel word-forms acquired through Hebb repetition learning of phonological sequences. One half of the participants were instructed to learn Hebb sequences of nonsense syllables in the morning (the morning group), whereas the other half learned the same materials in the evening (the evening group). Both groups completed three pause-detection tests, 0, 12 and 24 h after Hebb learning, in order to estimate whether the syllable sequences had created novel word-form entries in the mental lexicon. Because the second pause detection test was after a day of wakefulness in the morning group but after a night of regular sleep in the evening group, the procedure allowed us to estimate the relative contribution of sleep vs. passage of time to lexical engagement.

The pause detection results did not provide evidence for lexical competition immediately (0 h) after Hebb learning,

Table 3

Mean pause detection times (RT; ms) and accuracy (ACC; %correct) for base-words, control words and filler words as a function of delay after Hebb learning (0 h, 12 h and 24 h) in the morning group and evening group (Experiment 2). Standard deviations are in parentheses.

	Morning group			Evening group		
	0 h	12 h	24 h	0 h	12 h	24 h
RT						
Base	742 (137)	677 (132)	652 (122)	758 (169)	737 (151)	689 (116)
Control	750 (158)	650 (134)	625 (105)	760 (147)	707 (161)	656 (106)
Filler	603 (138)	522 (109)	513 (107)	638 (159)	566 (119)	528 (118)
ACC						
Base	93 (10)	92 (09)	92 (08)	91 (11)	91 (08)	91 (09)
Control	93 (13)	94 (06)	93 (08)	91 (13)	94 (09)	93 (10)
Filler	88 (09)	87 (07)	85 (05)	90 (07)	88 (06)	87 (07)

neither for the morning group, nor for the evening group. Twelve hours later (12 h) however, a lexical competition effect could be observed in both groups and it was still measurable at a similar magnitude 24 h after Hebb learning (24 h). These results support the notion of lexical consolidation by showing that learning novel word-forms does not immediately lead to a lexical representation but that lexical engagement clearly involves an incubation period. This period of lexical consolidation does, however, not necessarily require sleep. Clearly, the mere passage of time suffices to integrate the novel word memories in the mental lexicon and sleep does not especially accelerate this lexicalization process, at least not for lexical entries acquired through Hebb repetition learning.

Our findings do support a two-stage account of word learning (e.g., Davis et al., 2009) which proposes a functional separation between an initial stage of familiarization with the novel word-forms, followed by a slower offline integration of these word-forms in the mental lexicon. This initial familiarization phase has been hypothesized to rely on medial-temporal structures, like the hippocampus (e.g., Gooding, Mayes, & van Eijk, 2000), before slowly integrating the newly learnt information with existing knowledge, at the level of the neocortex (e.g., Tyler, Marslen-Wilson, & Stamatakis, 2005). The function of this intermediate hippocampal stage is to avoid catastrophic interference between the new and old knowledge (McClosky & Cohen, 1989). Whereas earlier studies have shown that the cortical changes in response to newly acquired word-forms are supported through or associated with sleep activity (Davis et al., 2009; Tamminen et al., 2010), our findings suggest that long-term lexical integration can also occur during wakefulness. Although sleep may offer optimal circumstances for offline lexical consolidation of newly acquired word-forms (for instance, because this implies less interference from other language input), some novel phonological representations, such as those established under Hebb repetition learning, gain access during wakefulness to the cortical networks that accommodate the mental lexicon. It is thus this gradual transfer of the new knowledge (novel phonological word-forms) from hippocampal to neocortical networks that is the critical mechanism operating during the passage of time. In summary, our study supports the idea of lexical consolidation within a two-stage account of word learning, but it falsifies the necessity of sleep to achieve lexical engagement. One interesting avenue for future research may be to investigate the characteristics of lexical consolidation in the absence of sleep.

General discussion

For several decades, memory theorists have tried to understand how exactly human memory supports the acquisition of novel words. Framed within the Page and Norris (2008, 2009) framework, the current Hebb-learning study is the first direct demonstration of how repetitive recall of phonological sequences gradually develops familiar word-forms that in turn, after an incubation period, are stored in the mental lexicon, where they acquire the status of a new word. This indicates that the cognitive processes responsible for representing serial order information in memory not only support short- and long-term serial recall of items (as in a span task or Hebb procedure), but also support the formation of the fundamental units of our language system, namely words. Our experimental findings are therefore a strong, empirical demonstration of how two major cognitive functions, namely memory and language, interact in the service of lexical learning.

As conceived within the Page and Norris (2008, 2009) framework, the current data show that, over repetitions, an ordered sequence of phonemes is likely to be learned as a chunk, with the chunk delimited by the temporal grouping structure and/or by the linguistic surroundings. In their model, the long-term learning of a chunk essentially comes down to the addition to memory of a new, localist representation that activates when its constituent items are presented in the correct serial order. Ultimately, the recall of the phonological items within a thoroughly learned chunk will be achieved via the activation of the single chunk representation, rather than by activation of the representations of the individual phonemes (see Page and Norris (2008, 2009) for implementational details). From this perspective, a word-form that is implicitly acquired through repetition-based learning is simply a chunked sequence of sublexical items that establishes stable, long-term representations in the mental lexicon over time. It is important to note that the present chunking account of word-form acquisition is congruent with the influential statistical learning approach taken by Saffran and colleagues (e.g., Saffran, Aslin, & Newport, 1996), who propose that word-forms are segmented from continuous speech based on statistical computations on probabilities of phoneme transitions (see also Fernandes et al., 2009). In this view, the chunking and statistical computation accounts are, although originating from different research traditions, both believed to describe the same domain-general learning mechanism (Perruchet & Pacton, 2006).

The present findings further show that the lexicalization of phonological sequences after Hebb learning requires some time, but also that this time interval does not necessarily need to contain sleep. On the one hand, this finding supports the two-stage account of word-form learning (Davis et al., 2009; McClelland, McNaughton, & O'Reilly, 1995; O'Reilly & Norman, 2002) which proposes that novel words are initially stored in the hippocampus, after which they are gradually transmitted to neocortical structures that are responsible for the long-term lexical representation of those novel words. On the other hand, our Hebb learning results do question the idea that long-term consolidation of novel words, and lexical interactions arising from these new representations, necessarily requires sleep (Davis et al., 2009; Dumay & Gaskell, 2007; Tamminen et al., 2010). We observed comparable lexical competition effects 12 h after Hebb learning, irrespective of whether our participants had slept or not, and the magnitude of this competition effect did not alter in the subsequent 12 h. The discrepancy between the Dumay and Gaskell (2007; Davis et al., 2009) findings and ours can presumably be attributed to the different paradigms that were used to make word learning operational, namely phoneme

monitoring and Hebb learning, respectively. Earlier in this paper, we compared both paradigms based on what Diekelmann et al. (2009) have identified as the determinants of sleep requirement for memory consolidation. Two factors appeared particularly relevant for the present study on lexical consolidation, namely the strength and explicitness of the lexical memory. Phoneme monitoring is a research task employed in the domain of auditory word recognition that is also used as a nonword exposure task in studies of language acquisition. Hebb repetition learning is a short-term serial recall task in which memory for one particular sequence is improved because it is frequently repeated throughout the experimental session. We presumed that Hebb learning establishes stronger memory traces than phoneme monitoring (learning vs. mere exposure) and the idea of deliberate word-form acquisition is more explicitly present in the phoneme monitoring than in the Hebb repetition procedure, in the instructions as well as in the format of the materials (auditory word-forms vs. visual syllable sequences). Hence, following Diekelmann et al. (2009), who argued that particularly weak and explicit memory traces are fed into the neocortex during sleep, the different learning characteristics of the Hebb vs. phoneme-monitoring paradigms offer a plausible explanation for the finding that lexical engagement of Hebb sequences does not necessarily require sleep.

Another recent study that is relevant in the light of these conclusions is the artificial language learning experiment by Fernandes et al. (2009). As we described earlier, Fernandes and her colleagues demonstrated that pseudowords that are segmented from continuous syllable streams, through regularities in transitional probabilities between the syllables, show lexical engagement just as novel word-forms do. Interestingly, in a subset of the experiments, Fernandes already observed markers of lexical engagement (i.e., lexical competition) immediately after artificial language learning. These findings are intriguing because not only do they minimize the role of sleep, they even undermine the two-stage account of word learning by showing full lexicalization without incubation. There may be different explanations for this finding: it could be related to the use of artificial language learning in the nonword familiarization phase or – more likely – to the use of lexical decision as a test of lexicalization. As the authors themselves acknowledge, there is some controversy about whether the lexical decision test is a genuine measure of lexicalization (Gaskell & Dumay, 2003; Goldinger, 1998; Wagenmakers et al., 2004), because it is more vulnerable to peripheral factors such as stimulus familiarity and/or strategic decision processes. It is precisely for this reason that lexical decision results should be interpreted with care and why researchers have put forward the pause-detection task (Mattys & Clark, 2002) as a useful alternative for assessing lexical engagement. The point is that pause detection "offers a measure of lexical activity in the absence of any explicit linguistic judgement – listeners are simply asked to monitor for periods of silence in speech" (Gaskell & Dumay, 2003, p. 118), for which reason it is less susceptible to strategic influences. We are admittedly not able to definitely resolve the question whether the online lexical consolidation observed by Fernandes et al. (2009) is due to the use of the lexical decision task. Nevertheless, we believe that the debate regarding the different measures of familiarization and lexicalization in relation to lexical learning is something that does deserve further attention. The framework put forward by Leach and Samuel (2007), in which lexical configural information about a word can be dissociated from its more dynamic lexical engagement, may turn out be helpful in this context. Also note that the lexical competition effects on pause detection, in this study as well as in earlier studies (Dumay & Gaskell, 2007; Gaskell & Dumay, 2003), do not emerge directly after familiarization with the novel word-forms but only after a period of offline consolidation (i.e., in line with a two-stage account of lexicalization). Knowing that familiarity is an instant and automatic consequence of activating an item in memory (e.g., Szmalec, Verbruggen, Vandierendonck, & Kemps, 2011), the finding that our lexical competition effects only emerge after several hours of offline consolidation thus strongly suggests that they are not simply driven by familiarity with the Hebb syllables. The pause-detection task, in combination with the lexical competition rationale, thus seems a valuable instrument for the assessment of whether and when a newly acquired word-form starts to behave like a "real word" (i.e., one that shows lexical engagement, Leach & Samuel, 2007).

Inherent to the use of the lexical competition approach (Gaskell & Dumay, 2003; see also Dumay & Gaskell, 2012) is the fact that most of the phonological content of the Hebb sequences (e.g., sa-fa-ra-...) is already represented in the mental lexicon (i.e., in the known word safari). This might raise the question whether the gradual Hebb learning seen in this study relies on chunking as much as does the classical Hebb procedure. First, it should be emphasized that our Hebb procedure does not simply expose word-forms to the participants. Participants observe individual syllables that are presented visually, one by one on a computer screen. We know of no alternative theoretical framework that can account for the transformation of the visual syllable sequence sa-fa-ra into the unitary phonological representation safara, while somehow avoiding chunking by virtue of prior lexical knowledge of the word safari. Second, if the Hebb learning of lexical competitors was substantially different from standard verbal Hebb learning, one might also predict that the two Hebb learning curves would show differences, possibly reflecting faster learning in the condition in which the Hebb sequences overlap with known words. In our 2009 study (Szmalec et al., 2009), participants were subjected to Hebb learning of nonsense syllables in the same way as in this study, the only difference being that the sequences (e.g., zi-lo-ka-...) did not overlap with (nor were neighbors of) existing words. Inspection of the Hebb learning curve in the Szmalec et al. (2009) study shows that filler and Hebb sequences begin at a performance of around 50–60% accurate, with the repetition learning resulting in final performance of around 80–90% accuracy for the Hebb sequences (filler sequences remain around 50–60%). As can be determined from Figs. 1 and 3, the pattern of results in the current study is very much in line with these earlier findings. Finally, the idea that phonological overlap with existing words does not notably alter the acquisition of novel word-forms

is at least consistent with the fact that two strongly overlapping word-forms can end up referring to totally different concepts (e.g., cap and cat).

Following the earlier working memory literature (Baddeley et al., 1998), the current study focuses on vocabulary acquisition as the introduction of a novel order-based phonological representation into the mental lexicon. In this view, sequential phonological forms are the key to vocabulary acquisition (e.g., Saffran et al., 1996). We are thus primarily interested in understanding how memory serves the development of such forms. Of course, the purpose of our language system is to communicate in a meaningful way, so we do acknowledge that besides lexical engagement, the mapping of novel word-forms to the semantic network is a major aspect of language learning. This mapping process is beyond the scope of this paper. Rather, the current Hebb learning paradigm is especially suited to studying the prerequisite for such lexico-semantic mapping, namely the creation of an order-sensitive lexical phonological representation.

Conclusion

The present research shows that the Hebb-learning paradigm can provide novel insights in a variety of aspects of human behavior where memory for serial order is involved, amongst which language learning and processing is probably the most striking example (e.g., Conway & Pisoni, 2008). It has allowed us to identify the memory processes that are involved in novel word-form learning (i.e. those responsible for the representation of serial order), and to question the role of overnight sleep in the lexical consolidation of those novel word-forms. Moreover, as can be derived from our recent findings on impaired Hebb learning in dyslexia (Szmalec, Loncke, Page, & Duyck, 2011), the Hebb paradigm might also have potential as a laboratory analogue of novel word learning in many other memory- and language-related research areas.

References

Baddeley, A. D., Gathercole, S. E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105,* 158–173.

Bowers, J. S., Davis, C. J., & Hanley, D. A. (2005). Interfering neighbors: The impact of novel word learning on the identification of visually similar words. *Cognition, 97,* 45–54.

Burgess, N., & Hitch, G. J. (2006). A revised model of short-term memory and long-term learning of verbal sequences. *Journal of Memory and Language, 55,* 627–652.

Conway, C. M., & Christiansen, M. H. (2001). Sequential learning in non-human primates. *Trends in Cognitive Sciences, 5,* 539–546.

Conway, C. M., & Pisoni, D. B. (2008). Neurocognitive basis of implicit learning of sequential structure and its relation to language processing. *Annals of the New York Academy of Sciences, 1145,* 113–131.

Davis, M. H., Di Betta, A., Macdonald, M. J. E., & Gaskell, M. G. (2009). Learning and consolidation of novel spoken words. *Journal of Cognitive Neuroscience, 21,* 803–820.

Diekelmann, S., Wilhelm, I., & Born, J. (2009). The whats and whens of sleep-dependent memory consolidation. *Sleep Medicine Reviews, 13,* 309–321.

Dumay, N., & Gaskell, M. G. (2005). Do words go to sleep? Exploring consolidation of spoken forms through direct and indirect measures. *Behavioral and Brain Sciences, 28,* 69–70.

Dumay, N., & Gaskell, M. G. (2007). Sleep-associated changes in the mental representation of spoken words. *Psychological Science, 18,* 35–39.

Dumay, N., & Gaskell, M. G. (2012). Overnight lexical consolidation revealed by speech segmentation. *Cognition, 123,* 119–132.

Duyck, W., Desmet, T., Verbeke, L., & Brysbaert, M. (2004). WordGen: A tool for word selection and nonword generation in Dutch, English, German, and French. *Behavior Research Methods Instruments and Computers, 36,* 488–499.

Duyck, W., Szmalec, A., Kemps, E., & Vandierendonck, A. (2003). Verbal working memory is involved in associative word learning unless visual codes are available. *Journal of Memory and Language, 48,* 527–541.

Fernandes, T., Kolinsky, R., & Ventura, P. (2009). The metamorphosis of the statistical segmentation output: Lexicalization during artificial language learning. *Cognition, 112,* 349–366.

Gaskell, M. G., & Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition, 89,* 105–132.

Gathercole, S. E., Hitch, G. J., Service, E., & Martin, A. J. (1997). Short-term memory and long-term learning in children. *Developmental Psychology, 33,* 966–979.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review, 105,* 251–279.

Gooding, P. A., Mayes, A. R., & van Eijk, R. (2000). A meta-analysis of indirect memory tests for novel material in organic amnesics. *Neuropsychologia, 38,* 666–676.

Gupta, P. (2003). Examining the relationship between word learning, nonword repetition, and immediate serial recall in adults. *Quarterly Journal of Experimental Psychology, 56,* 1213–1236.

Gupta, P. (2008). A computational model of nonword repetition, immediate serial recall, and nonword learning. In A. Thorn & M. Page (Eds.), *Interactions between short-term and long-term memory in the verbal domain.* Hove, UK: Psychology Press.

Hebb, D. (1961). Distinctive features of learning in the higher animal. In J. F. Delafresnaye (Ed.), *Brain mechanisms and learning* (pp. 37–46). Oxford, UK: Blackwell.

Jenkins, J. G., & Dallenbach, K. M. (1924). Oblivescence during sleep and waking. *American Journal of Psychology, 35,* 605–612.

Leach, L., & Samuel, A. G. (2007). Lexical configuration and lexical engagement: When adults learn new words. *Cognitive Psychology, 55,* 306–353.

Leclercq, A.-L., & Majerus, S. (2010). Serial-order short-term memory predicts vocabulary development: Evidence from a longitudinal study. *Developmental Psychology, 46,* 417–427.

Majerus, S., Poncelet, M., Elsen, B., & Van der Linden, M. (2006). Exploring the relationship between new word learning and short-term memory for serial order recall, item recall, and item recognition. *European Journal of Cognitive Psychology, 18,* 848–873.

Majerus, S., Poncelet, M., Greffe, C., & Van der Linden, M. (2006). Relations between vocabulary development and verbal short-term memory: The relative importance of short-term memory for serial order and item information. *Journal of Experimental Child Psychology, 93,* 95–119.

Majerus, S., Poncelet, M., Van der Linden, M., & Weekes, B. S. (2008). Lexical learning in bilingual adults: The relative importance of short-term memory for serial order and phonological knowledge. *Cognition, 107,* 395–419.

Mattys, S. L., & Clark, J. H. (2002). Lexical activity in speech processing: Evidence from pause detection. *Journal of Memory and Language, 47,* 343–359.

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

McClosky, M., & Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 109–165). New York: Academic Press.

Mosse, E. K., & Jarrold, C. (2008). Hebb learning, verbal short-term memory, and the acquisition of phonological forms in children. *The Quarterly Journal of Experimental Psychology, 61,* 505–514.

O'Reilly, R. C., & Norman, K. A. (2002). Hippocampal and neocortical contributions to memory: Advances in the complementary learning systems framework. *Trends in Cognitive Sciences, 6,* 505–510.

Page, M. P. A., Cumming, N., Norris, D., Hitch, G. J., & McNeil, A. M. (2006). Repetition learning in the immediate serial recall of visual and auditory materials. *Journal of Experimental Psychology: Learning Memory and Cognition, 32,* 716–733.

Page, M. P. A., & Norris, D. (1998). The primacy model: A new model of immediate serial recall. *Psychological Review, 105,* 761–781.

Page, M. P. A., & Norris, D. (2009). A model linking immediate serial recall, the Hebb repetition effect and the learning of phonological word-forms. *Philosophical Transactions of the Royal Society B, 364*, 3737–3753.

Page, M. P. A., & Norris, D. (2008). Is there a common mechanism underlying word-form learning and the Hebb repetition effect? Experimental data and a modelling framework. In A. Thorn & M. Page (Eds.), *Interactions between short-term and long-term memory in the verbal domain* (pp. 136–156). Hove, UK: Psychology Press.

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: Two approaches, one phenomenon. *Trends in Cognitive Science, 10*, 233–238.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 1926–1928.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language, 35*, 606–621.

Szmalec, A., Duyck, W., Vandierendonck, A., Barberá Mata, A., & Page, M. P. A. (2009). The Hebb repetition effect as a laboratory analogue of novel word learning. *Quarterly Journal of Experimental Psychology, 62*, 435–443.

Szmalec, A., Loncke, M., Page, M., & Duyck, W. (2011). Order or disorder? Impaired Hebb learning in dyslexia. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 1270–1279.

Szmalec, A., Verbruggen, F., Vandierendonck, A., & Kemps, E. (2011b). Control of interference during working memory updating. *Journal of Experimental Psychology: Human Perception and Performance, 37*, 137–151.

Tamminen, J., Payne, J. D., Stickgold, R., Wamsley, E. J., & Gaskell, M. G. (2010). Sleep spindle activity is associated with the integration of new memories and existing knowledge. *Journal of Neuroscience, 30*, 14356–14360.

Tyler, L. K., Marslen-Wilson, W. D., & Stamatakis, E. A. (2005). Differentiating lexical form, meaning, and structure in the neural language system. *Proceedings of the National Academy of Sciences, USA, 102*, 8375–8380.

Wagenmakers, E. J. M., Steyvers, M., Raaijmakers, J. G. W., Shiffrin, R. M., Van Rijn, H., & Zeelenberg, R. (2004). A model for evidence accumulation in the lexical decision task. *Cognitive Psychology, 48*, 332–367.