Psychological Science

Sleep-Associated Changes in the Mental Representation of Spoken Words

Nicolas Dumay and M. Gareth Gaskell

The online version of this article can be found at:

Published by:

$SAGE

http://www.sagepublications.com

On behalf of:

Association for Psychological Science

Additional services and information for *Psychological Science* can be found at:

Email Alerts: http://pss.sagepub.com/cgi/alerts

Subscriptions: http://pss.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Research Report

Sleep-Associated Changes in the Mental Representation of Spoken Words

Nicolas Dumay and M. Gareth Gaskell

University of York, York, United Kingdom

ABSTRACT—*The integration of a newly learned spoken word form with existing knowledge in the mental lexicon is characterized by the word form's ability to compete with similar-sounding entries during auditory word recognition. Here we show that although the mere acquisition of a spoken form is swift, its engagement in lexical competition requires an incubation-like period that is crucially associated with sleep. Words learned at 8 p.m. do not induce (inhibitory) competition effects immediately, but do so after a 12-hr interval including a night's sleep, and continue to induce such effects after 24 hr. In contrast, words learned at 8 a.m. do not show such effects immediately or after 12 hr of wakefulness, but show the effects only after 24 hr, after sleep has occurred. This time-course dissociation is best accommodated by connectionist and neural models of learning in which sleep provides an opportunity for hippocampal information to be fed into long-term neocortical memory.*

For models of human spoken-word recognition (cf. Jusczyk & Luce, 2002), a critical feature of any lexical entry is its ability to be evoked when compatible with the input, and to compete with similar-sounding entities for identification. Hence, an acid test of whether a newly learned spoken word form has been integrated into long-term lexical memory is whether it engages in this competition and inhibits recognition of phonological neighbors. In this article, we refer to the integration of a spoken form into lexical competition as "lexicalization" for short. In reality, this is only one component of the full lexicalization process, which also involves binding of form, syntax, and meaning.

A previous study established that—unlike phonological learning—lexicalization of a spoken form is relatively slow (Gaskell & Dumay, 2003). Through repeated exposure in phoneme monitoring, participants learned fictitious words that overlapped with existing words (e.g., "cathedr*uke*" for "cathedr*al*"). Using a two-alternative forced-choice (2-AFC) task (e.g., "cathedruke" vs. "cathedruce"), we showed good explicit recognition of novel words right after exposure, and no significant change 1 week later. The influence of novel words on lexical competition was assessed using pause detection (PD): Participants made speeded decisions about the presence of a silent pause inserted toward the offset of the existing words (e.g., "cathedr_al"). Mattys and Clark (2002; Mattys, Pleydell-Pearce, Melhorn, & Whitecross, 2005) showed that PD provides a good on-line measure of the overall level of competition and activity in lexical memory elicited by speech. For instance, PD latencies were shorter after early-unique words (which deviate from all other words well before their offset) than after late-unique words. Likewise, PD latencies correlated positively with the number of words compatible with the preceding speech. This suggests that activation of lexical candidates uses processing resources that would otherwise be allocated to PD. In view of these findings, we (Gaskell & Dumay, 2003) predicted that any increase in lexical activity induced by exposure to a novel competitor (e.g., "cathedruke") would produce longer PD latencies to a neighboring existing word (e.g., "cathedr_al"). We found that the negative impact on PD resulting from learning a new competitor was not observed immediately, but emerged during the time interval between exposure and retest, 7 days later. (A similar delay was found when the lexicalization test was lexical decision rather than PD.) Lexicalization therefore differs markedly from simple phonological (episodic) learning in that it requires an incubation-like period.

Address correspondence to Nicolas Dumay, Department of Experimental Psychology, University of Bristol, 12A, Priory Rd., Bristol BS8 1TU, United Kingdom, e-mail: n.dumay@bristol.ac.uk, or to M. Gareth Gaskell, e-mail: g.gaskell@psych.york.ac.uk.

Follow-up experiments (Dumay, Gaskell, & Feng, 2004) examined the time course of lexicalization more closely, tracking effects of exposure to novel words at three points: immediately after exposure, 24 hr later, and 7 days later. Again, there was no evidence of immediate lexicalization, but at both later time points, PD latencies demonstrated that the new competitor was contributing to lexical activity. Interestingly, performance on direct declarative measures of learning—2-AFC recognition and free recall—also improved over time. Furthermore, the same learning profile emerged irrespective of whether or not a meaning and a sentential context were supplied, suggesting that the lexicalization delay was not a by-product of impoverished encoding conditions. Most crucially, the observation of a change in lexical activity on the next day indicated that lexicalization of spoken word forms takes place within 24 hr after exposure.

The time-course difference between episodic learning and lexicalization is intriguing, and may reflect how humans overcome the catastrophic-interference problem faced by neural networks with distributed representations. In such networks, learning new information often disrupts preexisting knowledge (French, 1999). One computational solution to this problem relies on a dual-learning system in which new information is initially stored separately from existing knowledge and then interleaved off-line over a longer period of time (McClelland, McNaughton, & O'Reilly, 1995). According to some accounts, this interleaving process gradually transfers new information into long-term neocortical memory during sleep (French, 1999; Robins & McCallum, 1999; Wilson & McNaughton, 1994).

The role of sleep in memory consolidation has been documented mainly in the procedural domain (Maquet, 2001; Stickgold & Walker, 2005; Walker, 2005), and there is good evidence that sleep can improve perceptual and motor skills acquired during prior wakefulness (Karni, Tanne, Rubenstein, Askenasy, & Sagi, 1994; Walker, Brakefield, Morgan, Hobson, & Stickgold, 2002). Recent data on perceptual learning of speech indicate that sleep can also help recover procedural details acquired but gradually lost over the preceding day (Fenn, Nusbaum, & Margoliash, 2003), and can also promote abstraction of syntactic dependencies (Gómez, Bootzin, & Nadel, 2006).

The present study assessed whether nocturnal sleep (along with its physiological and neural correlates) enables listeners to lexicalize spoken words. Participants learned 24 nonsense competitor words (referred to as "novel" words) through phoneme monitoring either at 8 p.m. (p.m. group) or at 8 a.m. (a.m. group). Effects of exposure on lexical activity and explicit memory were measured immediately, 12 hr later, and 24 hr later. The first retest occurred after a night's sleep for the p.m. group but after a day awake for the a.m. group; conversely, the second retest occurred after a day awake for the p.m. group but after a night's sleep for the a.m. group. Changes in lexical activity were tracked using PD; explicit knowledge about the novel words was assessed using free recall and 2-AFC recognition (see Fig. 1a).
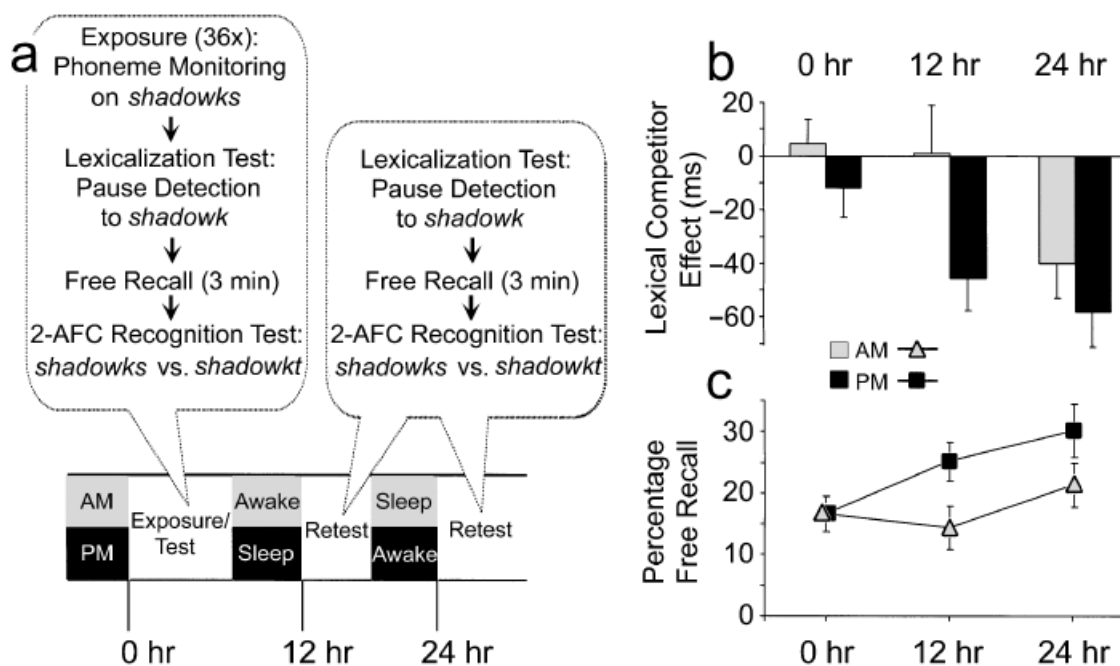

Fig. 1. Schematic representation of the experiment (a), mean lexical-competitor effect (response latency in the baseline condition minus response latency in the competitor condition) across pause-absent and pause-present trials (b), and correct response rate in free recall (c). Results are shown separately for the a.m. and p.m. groups. The two groups were exposed to the novel words in the morning or evening, respectively, and were tested immediately after exposure, 12 hr later, and after 24 hr. Error bars show standard errors.

METHOD

Participants

Sixty-four native English speakers (25 males; age range: 18–28; no known auditory, language, or sleep impairment) took part. The p.m. and a.m. groups were matched in terms of self-reported amount of sleep during the night interval of the experiment, 7 hr 24 min versus 7 hr 14 min, $t(62) < 1$.

Materials and Design

The key materials were 48 bisyllable triplets for which emergent competition effects had been demonstrated (Dumay et al., 2004). Each triplet included an English word (e.g., "shadow"; we refer to these words as the "existing" words) and two nonwords derived by adding final consonant clusters (e.g., "shadowks," "shadowkt"). One nonword was presented as a novel word, and the other was presented as a foil in the 2-AFC recognition test. Words were stress-initial morphologically simple nouns with an unreduced vowel as final segment and a uniqueness point before or on this segment.

During exposure, participants heard 24 novel words. Participants then heard all 48 of the existing words during the lexicalization test; half of these items potentially had a longer competitor as a result of exposure (the *competitor* condition), whereas the other half did not (the *baseline* condition). In both these conditions, half the existing words had a 200-ms pause inserted (see the Procedure section). We used four lists so that each item was equally represented in the four cells of the design (i.e., competitor, pause present; competitor, pause absent; baseline, pause present; baseline, pause absent). So we could assess the influence of the 2-AFC recognition test on the learning process, half the participants took this test only in the third session.

Procedure

Stimuli were delivered via headphones, using DMDX (Forster & Forster, 2003). The *exposure phase* (phoneme monitoring) involved 24 novel words, each played 36 times. On each trial, participants decided as quickly as possible whether a pre-specified phoneme was present.

In the *lexicalization test* (PD), participants made speeded decisions (by pressing one of two buttons) on whether a pause was present in each spoken stimulus. Each existing word was presented in a longer carrier compatible with the new competitor (e.g., "shadowk"). On pause-present trials, the pause was inserted a few cycles before the second vowel offset if the following consonant was a voiceless plosive and was inserted just after this vowel otherwise. Fillers were 144 bisyllabic words (72 with pauses inserted). Response latencies were measured from a marker representing pause onset.

In *free recall*, participants had 3 min to recall orally as many of the novel words as possible.

Finally, in the *2-AFC recognition test*, participants heard each novel word presented along with its foil and indicated (by button press) which of the two items was more familiar.

The first session (0 hr) comprised an exposure phase, followed by the lexicalization, the free-recall, and the 2-AFC recognition tests. The subsequent retests (12 and 24 hr) used the same test battery but involved no exposure phase (see Fig. 1a). Note that half the participants did not take the 2-AFC test until the third session.

RESULTS

To avoid confounds related to circadian differences in amount of initial acquisition, we analyzed performance of two subgroups of 24 participants matched on immediate performance in free recall (17%) and recognition (89%). (Analyses of the full data set showed an equivalent pattern.) Analyses of variance factored group, session, competitor acquisition (competitor vs. baseline condition), pause occurrence (presence vs. absence of a pause in PD), and inclusion of the 2-AFC recognition test in the session,[1] plus a dummy variable (Pollatsek & Well, 1995). PD latencies above 1,400 ms (2.7%) were excluded.

Performance in 2-AFC recognition indicated good knowledge of the novel words soon after exposure. Correct responses were near ceiling for all three sessions (88.2–90.3%), and there were no significant between-sessions variations for either group.

PD latencies revealed a clear association between sleep and the emergence of a change in lexical activity after acquisition of a new competitor (Fig. 1b). Exposure had no effect on the strength of lexical competition (i.e., the reaction time difference between the baseline and competitor conditions) in either group during the first session (Fs < 1). Items for which a new competitor had been learned were responded to no slower than those for which no competitor had been learned. After 12 hr, the p.m. group, who had experienced a night's sleep, showed a reliable competition effect (46 ms), $F_1(1, 32) = 7.56$, $p_{rep} = .97$, $\eta_p^2 = .52$; $F_2(1, 44) = 10.92$, $p_{rep} = .99$, $\eta_p^2 = .20$ (unless stated otherwise, all ps < .05; for information on calculating and interpreting p_{rep}, see Killeen, 2005), whereas the a.m. group, who had spent a day awake, did not show this effect (Fs < 1). Finally, 24 hr after exposure, the p.m. group continued to show a competition effect, $F_1(1, 32) = 21.84$, $p_{rep} = .998$, $\eta_p^2 = .59$; $F_2(1, 44) = 13.73$, $p_{rep} = .99$, $\eta_p^2 = .24$, and the a.m. group (having now slept) showed an equivalent effect, $F_1(1, 32) = 29.47$, $p_{rep} > .99$, $\eta_p^2 = .64$; $F_2(1, 44) = 6.05$, $p_{rep} = .95$, $\eta_p^2 = .12$.

Confirming the differential lexicalization pattern for the p.m. and a.m. groups, the three-way interaction of group, session, and competitor acquisition was significant for Sessions 2 and 3 combined, $F_1(1, 32) = 5.41$, $p_{rep} = .94$, $\eta_p^2 = .15$; $F_2(1, 44) = 4.34$, $p_{rep} = .92$, $\eta_p^2 = .09$. Also, there was no significant difference between the competition effects of the two groups before they went to sleep (Fs ≤ 1), or once they had slept (Fs ≤ 1.22). However, the competition effects showed reliable overnight increases in both groups: p.m. group—$F_1(1, 32) = 3.04$, $p = .09$,

---

[1]To avoid empty cells, we dropped this variable from the by-item PD analysis.

$p_{rep} = .88$, $\eta_p^2 = .22$; $F_2(1, 44) = 4.30$, $p_{rep} = .92$, $\eta_p^2 = .09$; a.m. group—$F_1(1, 32) = 12.46$, $p_{rep} = .99$, $\eta_p^2 = .33$; $F_2(1, 44) = 9.60$, $p_{rep} = .98$, $\eta_p^2 = .18$.

Although (episodic) learning was immediately evident on both explicit measures, free recall was also influenced by sleep, as the group-by-session interaction indicated, $F_1(2, 64) = 10.08$, $p_{rep} = .996$, $\eta_p^2 = .24$; $F_2(2, 88) = 13.12$, $p_{rep} = .999$, $\eta_p^2 = .23$ (Fig. 1c). Between Sessions 1 and 2, performance improved for the p.m. group (8.5%), $F_1(1, 32) = 24.65$, $p_{rep} > .99$, $\eta_p^2 = .62$; $F_2(1, 44) = 20.77$, $p_{rep} = .998$, $\eta_p^2 = .32$; but not for the a.m. group, which showed a trend toward poorer recall (−2.3%), $F_1(1, 32) = 2.36$, $p = .13$, $p_{rep} = .86$, $\eta_p^2 = .12$; $F_2(1, 44) = 2.84$, $p = .1$, $p_{rep} = .88$, $\eta_p^2 = .06$. In contrast, between Sessions 2 and 3, both groups improved—p.m. group: 5%; $F_1(1, 32) = 9.60$, $p_{rep} = .98$, $\eta_p^2 = .46$; $F_2(1, 44) = 11.00$, $p_{rep} = .99$, $\eta_p^2 = .20$; a.m. group: 7%; $F_1(1, 32) = 9.98$, $p_{rep} = .98$, $\eta_p^2 = .33$; $F_2(1, 44) = 12.52$, $p_{rep} = .99$, $\eta_p^2 = .22$. To assess whether reexposure during testing influenced performance, for half the participants we gave the clearest source of potential contamination—the 2-AFC test—only at the end of the experiment. This manipulation did not affect other measures. In any case, reexposure cannot explain the diverging trajectories of recall improvement and lexicalization for the p.m. and a.m. groups.

An account employing a fixed incubation period combined with differences in fatigue or cognitive efficiency at retest cannot explain our results either. First, if one supposed that greater fatigue or poorer efficiency in the evening had masked competition effects for the a.m. group when they were retested after 12 hr, then the competition effect should also have disappeared for the p.m. group at evening retest. Second, although the poorer recall in the evening than the morning for the a.m. group fits this explanation, the p.m. group showed best recall after 24 hr, again in the evening. Finally, existing data indicate that cognitive efficiency in young adults is in fact more optimal in the evening than in the morning. After typically 3 hr of sleep inertia, which manifests itself by a slow warm-up of neurocognitive functions, the detrimental effect of time awake is heavily counteracted throughout the day by a boost from the circadian pacemaker (Dijk, Duffy, & Czeisler, 1992; Johnson et al., 1992; Schneider, Fulda, & Schultz, 2004). In fact, our full data set also showed slightly better initial free recall for the p.m. group (19.4%) than the a.m. group (16.8%), $t_2(47) = 1.48$, $p = .072$, $p_{rep} = .85$, $d = 0.23$. If anything, therefore, an explanation involving circadian differences in fatigue or cognitive efficiency would predict, for both groups, competition effects and recall improvements in the evening retests only.

DISCUSSION

This study looked at the emergence of changes in the dynamics of lexical competition to examine whether sleep provides an opportunity for newly learned spoken forms to be incorporated in long-term lexical memory. The results support this hypothesis: Engagement in lexical competition was not observed shortly after exposure, nor after a day awake, but was consistently found after a night's sleep. The finding that the lexicalization of phonological forms is associated with nocturnal sleep (as opposed to the simple passing of time) is the primary contribution of our study.

As shown by free recall, declarative knowledge about the novel words was also improved after sleep. Whether this result reflects further strengthening of the episodic knowledge learned in the first place or is a consequence of the creation of new lexical entries remains an open question. Free recall may not tap exclusively into episodic memory, but could also rely on lexical knowledge when available. Therefore, although explicit measures can reflect only episodic knowledge prior to sleep (as lexicalization has not yet occurred), they may involve both episodic and lexical memory after sleep. In any case, the significant negative correlation between overnight improvement in recall and the postsleep competition effect in PD, $r(48) = −.35$, $p < .02$, indicates that enhancement of the new declarative knowledge and lexicalization were linked in some way (cf. Dumay & Gaskell, 2005). Participants showing strong competition after sleep also showed marked improvement in recall.[2] In contrast, there was no correlation between daytime changes in recall and postwake competition effects, $r(48) = .07$, n.s.

How does sleep enable lexicalization? Three potential interpretations of the findings are of interest. First, the key factor could involve some circadian state that is normally associated with sleep. In this case, whether the learner sleeps at night is immaterial; a brain state linked to the internal clock would be sufficient to trigger transfer to long-term memory. Second, although sleep does not prevent lexico-semantic stimulation and processing (Bastuji, Perrin, & Garcia-Larrea, 2002), it may be that merely the poverty of linguistic input (and thus interference) that normally characterizes sleep provides the necessary downtime for consolidation. Finally, it may be that sleep per se has a truly causal role, providing the neural or computational substrate necessary for lexicalization. Given that improvement in recall tasks is linked to specific configurations of non-REM/REM sleep cycles (Gais, Mölle, Helms, & Born, 2002; Plihal & Born, 1997), the correlation between the postsleep competition effect and the overnight improvement in recall strongly supports this possibility.

Sleep is known to enhance new procedural knowledge (Stickgold & Walker, 2005; Walker, 2005), whereas at first glance learning a word appears to be a classic declarative task. Our study demonstrates that declarative knowledge can also be enhanced after sleep. More important, though, we have shown that overnight sleep affects spoken-word recognition, a highly

---

[2]The correlation is negative because competition effects were calculated by subtracting response latencies in the (slower) competitor condition from baseline response latencies.

automatized skill. This dual consolidation effect of sleep may reflect the various types of knowledge encompassed by multi-faceted lexical representations.

In addition to linking phonological acquisition to sleep, our data provide support for distributed models of language and memory (Elman, 2004; McClelland et al., 1995; Robins & McCallum, 1999), which predict that a lexicalization delay is necessary to prevent new words from overwriting existing ones. Researchers have argued that such dual-learning systems make use of the hippocampus as a means of holding new information, which can then be fed progressively into the long-term neocortical store (McClelland et al., 1995; O'Reilly & Norman, 2002). Assuming that this characterization is correct, researchers can begin to associate aspects of vocabulary acquisition with the corresponding brain regions (Breitenstein et al., 2005). Hippocampal storage offers a (possibly temporary) basis for retrieval of specific information about newly acquired words, but information involving the phonological relationships between words, which are central to lexical competition, can reside only in the more integrated neocortical store.

REFERENCES

Bastuji, H., Perrin, F., & Garcia-Larrea, L. (2002). Semantic analysis of auditory input during sleep: Studies with event related potentials. *International Journal of Psychophysiology, 46*, 243–255.

Breitenstein, C., Jansen, A., Deppe, M., Foerster, A.F., Sommer, J., Wolbers, T., & Knecht, S. (2005). Hippocampus activity differentiates good from poor learners of a novel lexicon. *NeuroImage, 25*, 958–968.

Dijk, D.-J., Duffy, J.F., & Czeisler, C.A. (1992). Circadian and sleep/wake dependent aspects of subjective alertness and cognitive performance. *Journal of Sleep Research, 1*, 112–117.

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., & Feng, X. (2004). A day in the life of a spoken word. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the Twenty-Sixth Annual Conference of the Cognitive Science Society* (pp. 339–344). Mahwah, NJ: Erlbaum.

Elman, J.L. (2004). An alternative view of the mental lexicon. *Trends in Cognitive Sciences, 8*, 301–306.

Fenn, K.M., Nusbaum, H.C., & Margoliash, D. (2003). Consolidation during sleep of perceptual learning of spoken language. *Nature, 425*, 614–616.

Forster, K.I., & Forster, J.C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*, 116–124.

French, R.M. (1999). Catastrophic forgetting in connectionist networks. *Trends in Cognitive Sciences, 3*, 128–135.

Gais, S., Mölle, M., Helms, K., & Born, J. (2002). Learning-dependent increases in sleep spindle density. *Journal of Neuroscience, 22*, 6830–6834.

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

Gómez, R.L., Bootzin, R.R., & Nadel, L. (2006). Naps promote abstraction in language-learning infants. *Psychological Science, 17*, 670–674.

Johnson, M.P., Duffy, J.F., Dijk, D.-J., Ronda, J.M., Dyal, C.M., & Czeisler, C.A. (1992). Short-term memory, alertness and performance: A reappraisal of their relationship to body temperature. *Journal of Sleep Research, 1*, 24–29.

Jusczyk, P.W., & Luce, P.A. (2002). Speech perception and spoken word recognition: Past and present. *Ear and Hearing, 23*, 2–40.

Karni, A., Tanne, D., Rubenstein, B.S., Askenasy, J.J., & Sagi, D. (1994). Dependence on REM sleep of overnight improvement of a perceptual skill. *Science, 265*, 679–682.

Killeen, P.R. (2005). An alternative to null-hypothesis significance tests. *Psychological Science, 16*, 345–353.

Maquet, P. (2001). The role of sleep in learning and memory. *Science, 294*, 1048–1052.

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

Mattys, S.L., Pleydell-Pearce, C.W., Melhorn, J.F., & Whitecross, S.E. (2005). Detecting silent pauses in speech: A new tool for measuring on-line lexical and semantic processing. *Psychological Science, 16*, 958–964.

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.

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.

Plihal, W., & Born, J. (1997). Effects of early and late nocturnal sleep on declarative and procedural memory. *Journal of Cognitive Neuroscience, 9*, 534–547.

Pollatsek, A., & Well, A.D. (1995). On the use of counterbalanced designs in cognitive research: A suggestion for a better and more powerful analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 785–794.

Robins, A., & McCallum, S. (1999). The consolidation of learning during sleep: Comparing the pseudorehearsal and unlearning accounts. *Neural Networks, 12*, 1191–1206.

Schneider, C., Fulda, S., & Schultz, H. (2004). Daytime variation in performance and tiredness/sleepiness ratings in patients with insomnia, narcolepsy, sleep apnea and normal controls. *Journal of Sleep Research, 13*, 373–383.

Stickgold, R., & Walker, M.P. (2005). Memory consolidation and reconsolidation: What is the role of sleep? *Trends in Neurosciences, 28*, 408–415.

Walker, M.P. (2005). A refined model of sleep and the time course of memory formation. *Behavioral and Brain Sciences, 28*, 51–104.

Walker, M.P., Brakefield, T., Morgan, A., Hobson, J.A., & Stickgold, R. (2002). Practice with sleep makes perfect: Sleep dependent motor skill learning. *Neuron, 35*, 205–211.

Wilson, M.A., & McNaughton, B.L. (1994). Reactivation of hippocampal ensemble memories during sleep. *Science, 265*, 676–679.