Cognition 128 (2013) 280–286

Contents lists available at SciVerse ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Brief article

Sleep restores loss of generalized but not rote learning of synthetic speech

Kimberly M. Fenn [a,*], Daniel Margoliash [b,c,d], Howard C. Nusbaum [b,d]

[a] Department of Psychology, Michigan State University, East Lansing, MI 38824, United States
[b] Department of Psychology, University of Chicago, Chicago, IL, United States
[c] Department of Organismal Biology and Anatomy, University of Chicago, Chicago, IL, United States
[d] Center for Cognitive and Social Neuroscience, University of Chicago, Chicago, IL, United States

ARTICLE INFO

Article history:
Received 31 December 2010
Revised 26 March 2013
Accepted 22 April 2013
Available online 4 June 2013

Keywords:
Sleep
Memory
Consolidation

ABSTRACT

Sleep-dependent consolidation has been demonstrated for declarative and procedural memory but few theories of consolidation distinguish between rote and generalized learning, suggesting similar consolidation should occur for both. However, studies using rote and generalized learning have suggested different patterns of consolidation may occur, although different tasks have been used across studies. Here we directly compared consolidation of rote and generalized learning using a single speech identification task. Training on a large set of novel stimuli resulted in substantial generalized learning, and sleep restored performance that had degraded after 12 waking hours. Training on a small set of repeated stimuli primarily resulted in rote learning and performance also degraded after 12 waking hours but was not restored by sleep. Moreover performance was significantly worse 24-h after rote training. Our results suggest a functional dissociation between the mechanisms of consolidation for rote and generalized learning which has broad implications for memory models.

1. Introduction

The acquisition of complex skills depends on the ability to generalize beyond exact situations experienced during learning. It has been argued that the ability to generalize is the defining feature of adaptive learning, and the quality that distinguishes it from simple associative memory (Poggio & Bizzi, 2004). Many models of memory suggest that generalized learning relies on the same underlying associative mechanisms as learning of specific experiences; generalization depends on abstraction from associations acquired during training (e.g., Goldinger, 1998; Hintzman, 1986; McClelland & Rumelhart, 1985). In contrast, other theories suggest that memory involves both specific representations and abstract representations (cf. Anderson et al., 2004; Grossberg, 1986; Posner & Keele, 1968). Evidence suggesting that there may be different mechanisms underlying rote and generalized learning would present a challenge for models that posit only specific representations and would provide support for models that allow for both specific and abstract representations. Here we report that the two forms of learning show different patterns of sleep-dependent consolidation.

Memory consolidation research suggests that sleep consolidates procedural and perceptual skills (see Margoliash & Fenn, 2008; McGaugh, 2000; Walker, 2005, for reviews) but the vast majority of this research has emphasized tasks wherein learning is restricted to the exact information encountered during training. Tasks used to study procedural consolidation typically focus on learning one motor pattern (cf. Fischer, Hallschmid, Elsner, & Born, 2002) or discrimination of one visual pattern (cf. Karni, Tanne, Rubenstein, Askenasy, & Sagi, 1994) which may be considered rote procedural learning. In rote motor learning, sleep is reported to enhance learning; performance is

* Corresponding author. Tel.: +1 517 432 6258; fax: +1 517 432 2476.
E-mail address: kfenn@msu.edu (K.M. Fenn).

0010-0277/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cognition.2013.04.007

significantly better after sleep than after training, an effect not seen after an equal interval of wakefulness (cf. Fischer et al., 2002; Walker, Brakefield, Hobson, & Stickgold, 2003). Although several studies have reported this effect, recent work has argued that apparent memory enhancements may be explained by reactive inhibition (Rickard, Cai, Rieth, Jones, & Ard, 2008) or circadian differences at test (Cai & Rickard, 2009).

Although some of the work in rote procedural learning has been questioned, there is strong evidence that sleep consolidates generalized learning and promotes abstraction of information. In generalized procedural learning, performance degrades across waking retention and is restored by sleep. Sleep also inoculates memory against subsequent degradation (Brawn, Fenn, Margoliash, & Nusbaum, 2008; Fenn, Nusbaum, & Margoliash, 2003). Consistent with this, we have reported that after controlling for reactive inhibition, rote motor learning follows the same general pattern of waking degradation and restoration after sleep (Brawn, Fenn, Nusbaum, & Margoliash, 2010). Other studies that have investigated generalized learning and sleep have shown that sleep can restructure information acquired during waking. Of note, Wagner, Gais, Haider, Verleger, and Born (2004) trained participants on a complex algorithm that contained a hidden rule that allowed the problem to be solved in fewer steps. Participants were more likely to become aware of the hidden rule if tested after sleep than after a waking interval. Similarly, infants who were exposed to an artificial language showed evidence of generalization and abstraction of the rules of the language after a nap. In contrast, after a waking interval, infants showed stronger veridical memory, but did not show any evidence of abstraction or generalization (Gomez, Bootzin, & Nadel, 2006). Thus, there is strong evidence that sleep consolidates generalized learning and promotes abstraction or restructuring of information.

Given that sleep consolidates both rote and generalized learning, the potential difference in consolidation of these types of learning can be used to investigate whether different mechanisms underlie these two forms of learning. Consolidation in rote skills may be confined to lower-level cortices (cf. Karni & Bertini, 1997) whereas generalized learning may depend on the interaction of broader networks of neural activity (Ahissar & Hochstein, 2004; Poggio & Bizzi, 2004). Generalized skills may receive a different benefit from consolidation processes and may be more susceptible to waking interference.

The effects of sleep on rote and generalized learning have only been compared across substantially different tasks, complicating interpretation of differences. We compared rote and generalized learning in a synthetic speech learning task and tested the effect of waking retention and sleep on performance.

2. Method

2.1. Participants

We recruited 67 right-handed native English speakers who had no history of speech, hearing, or memory disorders. Nine participants were excluded from all analyses for not being native English speakers ($n = 4$), or for consuming alcohol on the study evening ($n = 1$), or for not completing the experiment ($n = 4$). The remaining 58 participants (32 female) had a mean age of 20.6 ± 3.6 (s.d.) years. All were students or employees at the University of Chicago and were financially compensated.

2.2. Materials

Seven hundred monosyllabic words were generated by *rsynth*, a text-to-speech synthesizer based on Klatt (1980). The intelligibility of this synthetic speech is relatively low, but listeners show significant improvement after one training session (Fenn et al., 2003). The words were taken from phonetically balanced lists, approximating the distribution of phonemes in English (Egan, 1948). Words were chosen based on the distribution of phonetic properties in English and were derived from a diverse set of syntactic categories (nouns, verbs, adjectives, and adverbs). The words were divided into four (100-word) test sets and one (300-word) training set. The test sets were balanced to establish comparable difficulty, based on pilot testing. Participants received each of the four tests in one of the following positions: Pretest, Posttest I, Posttest II, and Posttest III. Test order was counterbalanced across participants. In addition, twenty of the training words were added to each test so each test included 100 novel words and 20 words that repeated throughout the experiment (i.e. during training and every test).

2.3. Design

All participants completed three experimental sessions. The first session contained a pretest, training, and posttest. The second and third sessions contained only a posttest. Each session was separated by a 12-h retention interval. The first session was conducted at 09:00; the second session was at 21:00 (after waking retention), and the third session was at 09:00, after a retention interval that included sleep (see Supplementary Fig. 1). We did not test for circadian differences because our previous research demonstrated that time of day effects are negligible (Fenn et al., 2003). Participants were randomly assigned to one of two training procedures that have been found to produce rote or generalized learning respectively (Greenspan, Nusbaum, & Pisoni, 1988). One group (rote-trained) was trained on 20 words. Words were presented in pseudorandom order 15 times throughout training. The other group (generalization-trained) was trained on 300 unique stimuli. Twenty of the training stimuli for this group were used as the full training set for the rote-trained group and were used in each of the tests. The remaining 280 were novel words that did not appear in any of the tests.

An additional control group ($n = 12$) was trained exactly as the rote-trained group but was tested only on the 20 repeated stimuli during each posttest, to reduce interference during the tests and to ensure that performance in the rote-trained experimental group was not affected by the inclusion of novel items during testing. Because the rote-trained group experienced only 20 stimuli during training but encountered an additional 100 unique stimuli at test, these stimuli may have interfered with the consolidation process (cf. Nader, Schafe, & LeDoux, 2000).

2.4. Procedure

The pretest and each of the posttests for the rote-trained and generalization-trained experimental groups contained 120 stimuli: 100 novel words and the 20 repeated words (Supplementary Fig. 1). For the rote control group, each posttest contained only the 20 repeated words. Individual stimuli were randomly presented via Sennheiser HD 570 headphones and participants identified stimuli by typing the word into the computer. There was no time limit imposed on responses.

During training, participants first received perceptual training on 50 words and were then tested on 20 words. This sequence (perceptual training and test) was repeated six times throughout training for a total of 300 perceptual training trials and 120 test trials. During perceptual training, individual words were presented via headphones, paired with the orthographic form of the word on the computer screen. The orthographic form preceded the auditory stimulus so participants could use stimulus identity to help shape perception. The generalization-trained group received perceptual training on 50 unique stimuli in each training set. The rote-trained and control groups received perceptual training on 50 stimuli during each set, but because the training set contained only 20 items, some words were repeated during each training phase. Across all training trials, each word was repeated 15 times. However, stimuli were delivered randomly across training blocks. Therefore, it was possible for one word to appear several times in a given training block and a different word to not appear at all during a given block. After 50 perceptual training trials, participants were tested on 20 words. For the generalization-trained group, the words were randomly chosen from the 50 they just heard. The rote-trained and rote control groups were tested on the full set of 20 training words, regardless of which items they experienced in the preceding perceptual training trials. During each test, participants had 4 s to complete a response. If a response was not completed in this time, the trial was scored as incorrect and the next stimulus was presented.

At the start of each experimental session, participants completed the Stanford Sleepiness Scale (Hoddes, Zarcone, Smythe, Phillips, & Dement, 1973) to assess sleepiness and visual analog mood scale (Monk, 1989) which measures global vigor and global affect (Supplemental online material).

3. Results

To assess immediate, practice-dependent improvements, we compared Pretest and Posttest I performance for novel and repeated words for the experimental groups. Performance was calculated as percent correct identification. Stimulus identification was scored as correct if the correct phonetic content of the word was identified. Incorrect spelling was not penalized (e.g. the word "kar" was scored correct if the stimulus was "car"). We first assessed performance on novel words by comparing identification on Pretest and Posttest I for the 100 unique words in each test. Performance on novel words improved for both groups, but the generalization-trained group showed larger improvements than rote-trained (17.5 ± 1.4 and 4.0 ± 1.6 mean percentage points ± S.E.M., respectively, Fig. 1a). Using a repeated measures ANOVA with training condition (rote, generalized) as a between-subjects factor and test (Pretest, Posttest I) as a within-subjects factor, we found a main effect of test ($F_{1,44} = 102.76$, $p < .001$), training condition ($F_{1,44} = 11.9$, $p < .01$), and an interaction between factors ($F_{1,44} = 40.21$, $p < .001$). Although the rote-trained group's improvement on novel words was modest, it was significant, $t(23) = 2.53$, $p < .01$.

Both groups also showed improvement on the 20 repeated words ($F_{1,44} = 526.65$, $p < .001$) (Fig. 1b), but the pattern of improvement was opposite to that of novel words. Rote-trained participants improved significantly more (51 ± 1.9) than generalization-trained participants (22.5 ± 2.6), as suggested by a significant effect of training

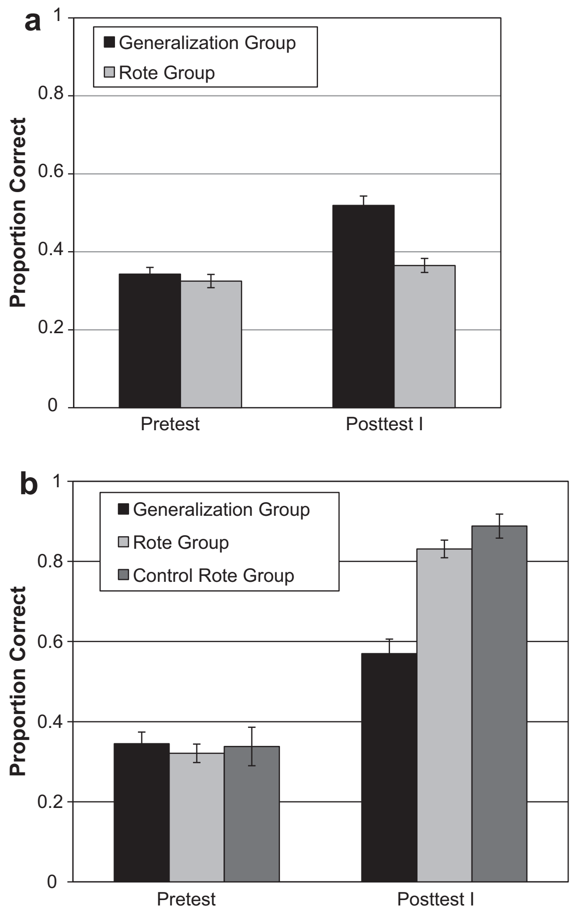

Fig. 1. Pretest and Posttest I performance, in proportion correct (±S.E.M.), for experimental groups and control groups. (a) Performance on novel words for the generalization-trained and rote-trained experimental groups and (b) performance on the 20 words used for training in the rote-trained group for the generalization-trained and rote-trained experimental groups and the rote-trained control group.

condition ($F_{1,44} = 10.99$, $p < .01$) and interaction between training and test ($F_{1,44} = 79.32$, $p < .001$).

Performance during each of the 20-item tests during training was scored in the same way as the tests. Fig. 2 shows that performance improved throughout training ($F_{5,220} = 25.92$, $p < .0001$), and that performance was better for the rote-trained group than the generalization-trained group ($F_{1,220} = 14.03$, $p < .001$). The interaction between test point and training condition was not significant ($F_{5,220} = 1.64$, $p = .15$).

The primary question of this study was to investigate the change in performance across time and sleep. To measure the effect of the two 12-h retention intervals, we compared performance across the three posttests. We first normalized scores for baseline differences by subtracting pretest accuracy from posttest accuracy (separately for repeated and novel items) and performed analyses on these improvement scores.

3.1. Novel items

Improvement on novel items was analyzed using a repeated measures ANOVA with training condition as a between-subjects factor and posttest as a within-subjects factor. We found a significant effect of training condition ($F_{1,44} = 25.49$, $p < .001$), posttest ($F_{2,44} = 10.71$, $p < .001$), and a significant interaction between the factors ($F_{2,88} = 7.29$, $p < .01$). Planned comparisons show that the generalization-trained group replicated our previous studies (Fig. 3a); performance was significantly worse after waking retention (Posttest II) than immediate test (Posttest I) $t(21) = 4.72$, $p < 0.001$ (all t-tests are one-tailed,[1] except where noted) and performance significantly improved again (by $7 \pm 1.8\%$ points) after sleep at Posttest III (compared to Posttest II) $t(21) = 3.67$, $p < .001$. A small set of participants in this group ($n = 4$) were erroneously given the same 100-item test at two points during the experiment. Removing these participants did not affect the results; performance was significantly worse after waking retention $t(17) = 5.48$, $p < .001$ and significantly improved again after sleep $t(17) = 3.89$, $p < .001$.

Fig. 3a shows that the rote-trained group did not show a significant loss of performance on novel stimuli after waking retention $t(23) = .33$, $p = .37$, likely because the immediate improvement on novel test items was so small that there was little sensitivity to measure performance reduction. This group showed a trend for improved performance (compared to Posttest II) after sleep ($2.6 \pm 1.4\%$ points) that narrowly missed significance $t(23) = 1.48$, $p = .07$. Three participants in this group were also tested erroneously with the same test during two tests and therefore did not get a true generalization test on novel words. If we exclude these participants, the comparison between immediate and 12-h testing is no different, $t(20) = .29$, $p = .39$ but this group shows significantly improved performance after sleep ($3.0 \pm 1.6\%$ points), compared to Posttest II $t(20) = 1.74$, $p = .04$, suggesting that there is some consolidation of generalized learning for both groups.

[1] One-tailed t-tests were used because in three previous studies, we found a consistent pattern of performance across 12-h delay intervals. When training occurred in the morning, performance decreased across a waking interval and was restored after sleep (Brawn et al., 2008; Brawn et al., 2010; Fenn et al., 2003). Based on this work, we had a priori predictions for performance change across these delay intervals.

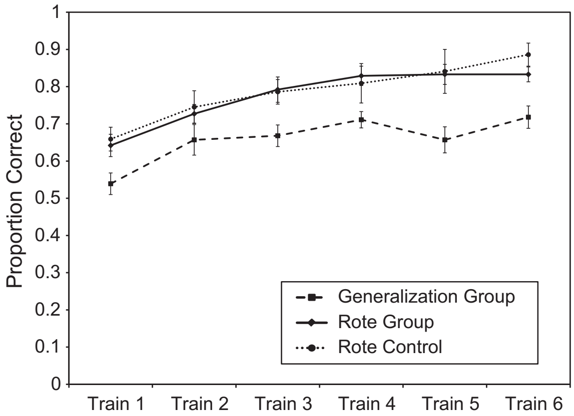

Fig. 2. Proportion correct (± S.E.M.) on each of the six tests during training for the three groups. All groups were tested on 20 items. For the rote-trained experimental group and the rote control group, the twenty items were the full training set; for the generalization-trained group the 20 items were randomly chosen from the 50 stimuli on which they just received perceptual training.

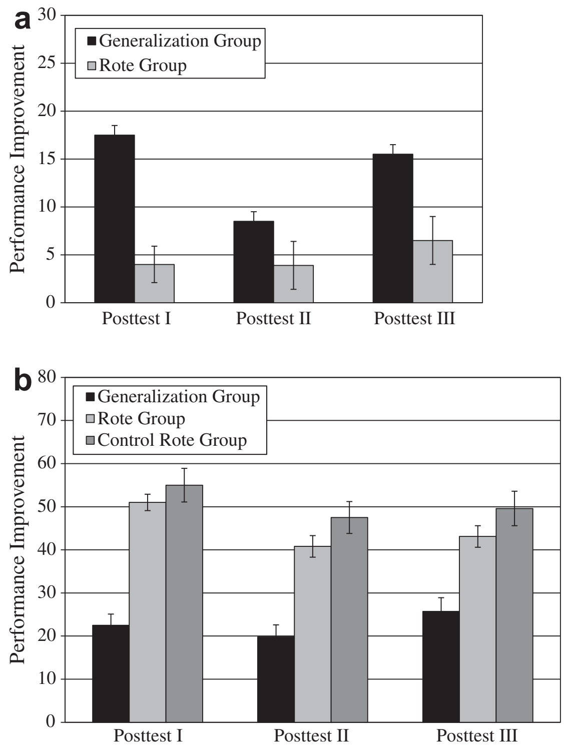

Fig. 3. Improvement (± S.E.M.) measured by change in performance (percent correct) from pretest to posttest for each of the three posttests: given immediately after training (Posttest I), 12-h after training (Posttest II), and 24-h after training (Posttest III). (a) Performance on novel words for the two experimental groups and (b) performance on repeated words for the two experimental groups and the control group. *Note*: graphs are not on the same scale.

3.2. Repeated items

We first analyzed performance on the repeated items for the two experimental groups. The pattern of results on the 20 repeated words was similar to the results on novel stimuli for the generalization-trained group (Fig. 3b), but was quite different for the rote-trained group. Using a repeated-measures ANOVA with training condition as a between-subjects factor and posttest as a within-subjects factor, we found a significant effect of training condition ($F_{1,44} = 45.83$, $p < .0001$), posttest ($F_{2,44} = 11.13$, $p < .001$), and a significant interaction between the factors ($F_{2,88} = 8.32$, $p < .001$). The generalization-trained group showed a trend for a reduction in performance following waking retention but this loss did not reach significance $t(21) = 1.42$, $p = .08$, possibly because the stimuli were no longer novel, having been heard during training and on two previous tests. Although these participants did not show significant reduction after waking, they did show improvement from Posttest II, of 5.9 ± 1.8% points, after sleep, $t(21) = 3.09$, $p < .01$.

Perhaps the most interesting pattern of results is the performance of the rote-trained group on the 20 repeated items. Similar to our previous studies, performance was significantly reduced after 12 waking hours $t(23) = 5.26$, $p < .001$. Interestingly, after sleep, this group did not show the expected rebound in performance. This group improved from Posttest II by an average of only 2.3 ± 1.9% points, a change that was not statistically reliable $t(23) = 1.31$, $p = .11$, although the direction of the effect is for increased performance. This may reflect some consolidation of generalization learning from the rote training rather than an effect of sleep on rote memory. Furthermore, this group showed significantly worse performance at the 24-h test than at immediate test $t(23) = 3.94$, $p < .001$ (two-tailed), an effect not previously observed. In contrast, on the same stimuli, the generalization group showed a trend for *better* performance on the 24-h test than on the immediate test, but this effect did not reach significance $t(21) = 1.68$, $p = .11$ (two-tailed).

The control rote group was tested only on the 20 repeated items to ensure that performance in the rote-trained experimental group (tested on 120 items on each test) was not affected by the inclusion of novel items during testing. The results from the rote control group closely resembled the results in the rote-trained experimental group (Fig. 3b). The control group showed robust improvement directly after training and a repeated-measures ANOVA found that improvement varied by test ($F_{2,22} = 6.52$, $p < .01$). Planned comparisons showed that performance was significantly worse after waking $t(11) = 3.35$, $p < .01$ and the change in performance after sleep was not significant $t(11) = .94$, $p = .18$, although there was a trend for increased performance. Again, performance was significantly worse on the 24-h test than at immediate posttest $t(11) = 2.41$, $p = .03$ (two-tailed). In this task, rote perceptual learning either did not benefit from sleep, or benefited far less than generalized learning.

Finally, previous research has shown that weak declarative memory obtains a greater benefit from sleep-related consolidation processes than stronger memory (Drosopoulos, Schultze, Fischer, & Born, 2007). We investigated the relationship between initial learning and subsequent consolidation in our groups. If sleep preferentially consolidated weaker associations in this task, then we would expect that the change in performance across sleep would be negatively correlated with either initial learning or performance on the first posttest. However, there was no evidence for a negative relationship in either group. The correlation between initial learning (change in performance from Pretest to Posttest I) and consolidation (change from Post II to Post III), in the rote-trained group was near zero on rote stimuli $r = .009$, $z = .04$, $p = .96$ and novel stimuli, $r = .18$, $z = .84$, $p = .39$. Similarly, if absolute performance on the first posttest was used as the measure of learning, the correlation between performance and consolidation was near zero for rote $r = .06$, $z = .31$, $p = .75$ and novel stimuli $r = .08$, $z = .36$, $p = .71$. For the generalization-trained group, there was not a significant relationship between initial learning and consolidation $r = .26$, $z = 1.16$, $p = .24$ or Posttest I performance and consolidation on rote stimuli $r = .25$, $z = 1.11$, $p = .26$. However, on novel stimuli, there was a marginally significant *positive* correlation between initial learning and consolidation $r = .38$, $z = 1.78$, $p = .07$ and Posttest I performance and consolidation $r = .41$, $z = 1.91$, $p = .055$. Thus, stronger initial learning was associated with increased consolidation in this group.

4. Discussion

The present study demonstrates one fundamental difference between rote and generalized learning. Sleep consolidated generalized learning but by comparison, rote memorization of 20 specific patterns did not evince any benefit from sleep-related consolidation processes. Following generalization training, recognition performance significantly degraded across waking and was restored after sleep. After rote training, performance diminished over waking retention, was not restored by sleep, and was significantly worse 24-h after training than at immediate test. Although rote training received very little benefit from sleep on items encountered during training, there was a trend for a rebound effect in generalization to novel items. Rote-trained listeners showed some (albeit much smaller) generalization from their training. This smaller generalization effect showed a pattern after sleep that was similar to the pattern displayed by the generalization-trained group, suggesting there might be consolidation of the generalization learning although not of the rote learning even within this one group of learners.

These results suggest that there may be different mechanisms underlying rote and generalized learning and challenge some views of memory processing. For learning to generalize and apply to novel contexts, it is important to abstract pattern information such that it can be applied to new information. Theories of memory address generalization in different ways. One approach is to represent only the specific experiences encoded during learning and to abstract from those experiences (e.g., McClelland & Rumelhart, 1985). In these models, generalization depends on rote memory traces from specific experiences (cf. Goldinger, 1998). Consolidation of generalized learning would therefore depend on consolidation of individual traces acquired during training as they form the memory base supporting generalization. The present data argue against these models because rote and generalized learning did not show the same benefit of consolidation.

An alternative class of models argues for both representational specificity and abstraction (cf. Posner & Keele, 1968). For example, speech experiments involving shadowing or priming have been used to argue that listeners' memory for speech contains both abstract representations and talker-specific acoustic patterns (e.g., Ju & Luce, 2006; McLennan, Luce, & Charles-Luce, 2005). Theories of memory as diverse as ART (Grossberg, 1986) and ACT-R (Anderson et al., 2004) provide for both abstract and specific representations. While these models do not address consolidation during sleep, the existence of specific and general representations, and the differentiation of types of representation and processing provide the separation that is necessary for the different effects of consolidation we report here. Consistent with this, several studies have shown that learning of new vocabulary words occurs quite rapidly. However, integration of new words into the lexicon requires time (Davis, Di Betta, Macdonald, & Gaskell, 2008) and, importantly, sleep (Dumay & Gaskell, 2007).

We propose that learning induces abstract representations based on the statistical distribution of experiences and that sleep consolidates these abstract representations. Recent evidence that sleep modifies daytime activity properties of neurons supports the idea that new information is created during sleep (Dragoi & Tonegawa, 2011; Rauske, Chi, Dave, & Margoliash, 2010). Initially, during a waking period, memory of the newly-induced abstraction degrades, as does memory of the specific tokens that helped form the abstract representation. During sleep, associative interactions through neuronal replay help to stabilize memory of the abstract representation and exemplar tokens. Additionally, an interaction between memory of tokens and abstract representations helps to stabilize both representations. For rote learning, there is little distributional information available for the induction of abstraction but considerable information for specific tokens. For generalized learning, there is more information in memory for the abstraction, and the interaction between the two is stronger. Such differences could account for the differences in behavior observed for the two types of learning and suggest a functional dissociation in the mechanisms underlying the consolidation of these two types of learning.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.cognition.2013.04.007.

References

Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences, 8*(10), 457–464.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review, 111*(4), 1036–1060.

Brawn, T. P., Fenn, K. M., Margoliash, D., & Nusbaum, H. C. (2008). Consolidation of sensorimotor learning during sleep. *Learning and Memory, 15*, 815–819.

Brawn, T. P., Fenn, K. M., Nusbaum, H. C., & Margoliash, D. (2010). Consolidating the effects of waking and sleep on motor-sequence learning. *Journal of Neuroscience, 30*(42), 13977–13982.

Cai, D. J., & Rickard, T. C. (2009). Reconsidering the role of sleep for motor memory. *Behavioral Neuroscience, 123*(6), 1153–1157.

Davis, M. H., Di Betta, A. M., Macdonald, M. J. E., & Gaskell, M. G. (2008). Learning and consolidation of novel spoken words. *Journal of Cognitive Neuroscience, 21*(4), 803–820.

Dragoi, G., & Tonegawa, S. (2011). Preplay of future place cell sequences by hippocampal cellular assemblies. *Nature, 469*(7330), 397–401.

Drosopoulos, S., Schultze, C., Fischer, S., & Born, J. (2007). Sleep's function in the spontaneous recovery and consolidation of memories. *Journal of Experimental Psychology: General, 136*(2), 169–183.

Dumay, N., & Gaskell, M. G. (2007). Sleep-associated changes in the mental representation of spoken words. *Psychological Science, 18*(1), 35–39.

Egan, J. P. (1948). Articulation testing methods. *Laryngoscope, 58*, 955–991.

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

Fischer, S., Hallschmid, M., Elsner, A. L., & Born, J. (2002). Sleep forms memory for finger skills. *Proceedings of the National Academy of Sciences of the United States of America, 101*, 2140–2144.

Goldinger, S. D. (1998). Echo of echoes? An episodic theory of lexical access. *Psychological Review, 105*(2), 251–279.

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

Greenspan, S. L., Nusbaum, H. C., & Pisoni, D. B. (1988). Perceptual learning of synthetic speech produced by rule. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(3), 421–433.

Grossberg, S. D. (1986). The adaptive self-organization of serial order in behavior: Speech, language, and motor control. In E. C. Schwab & H. C. Nusbaum (Eds.), *Pattern recognition by humans and machines. Speech perception* (Vol. 1, pp. 187–294). New York: Academic Press.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace model. *Psychological Review, 93*, 411–428.

Hoddes, E., Zarcone, V., Smythe, H., Phillips, R., & Dement, W. C. (1973). Quantification of sleepiness: A new approach. *Psychophysiology, 10*(4), 431–436.

Ju, M., & Luce, P. A. (2006). Representational specificity of within-category phonetic variation in the long-term mental lexicon. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 120–138.

Karni, A., & Bertini, G. (1997). Learning perceptual skills: Behavioral probes into adult cortical plasticity. *Current Opinion in Neurobiology, 7*, 530–535.

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

Klatt, D. H. (1980). Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America, 67*, 971–995.

Margoliash, D., & Fenn, K. M (2008). Sleep and memory consolidation in audition. In P. Dallos & D. Oertel (Eds.). *The senses: A comprehensive reference audition* (Vol. 3, pp. 895–912). San Diego: Academic Press.

McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General, 114*, 159–197.

McGaugh, J. L. (2000). Memory – A century of consolidation. *Science, 287*, 248–251.

McLennan, C. T., Luce, P. A., & Charles-Luce, J. (2005). Representation of lexical form: Evidence from studies of sublexical ambiguity. *Journal of Experimental Psychology: Human Perception and Performance, 31*, 1308–1314.

Monk, T. H. (1989). A visual analogue scale technique to measure global vigor and affect. *Psychiatry Research, 27*, 89–99.

Nader, K., Schafe, G., & LeDoux, J. (2000). Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. *Nature, 406*, 722–726.

Poggio, T., & Bizzi, E. (2004). Generalization in vision and motor control. *Nature, 431*, 768–774.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology, 77*, 353–363.

Rauske, P. L., Chi, Z., Dave, A. S., & Margoliash, D. (2010). Neuronal stability and drift across periods of sleep: Premotor activity patterns in a vocal control nucleus of adult zebra finches. *Journal of Neuroscience, 30*(7), 2783–2794.

Rickard, T. C., Cai, D. J., Rieth, C. A., Jones, J., & Ard, M. C. (2008). Sleep does not enhance motor sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(4), 834–842.

Wagner, U., Gais, S., Haider, H., Verleger, R., & Born, J. (2004). Sleep inspires insight. *Nature, 427*, 352–355.

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., Hobson, J. A., & Stickgold, R. (2003). Dissociable stages of human memory consolidation and reconsolidation. *Nature, 425*, 616–620.