THE ROLE OF SLEEP IN RETENTION AND GENERALIZATION OF STATISTICALLY LEARNED LANGUAGE

by

Lucia Sweeney

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF PSYCHOLOGY

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2020

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by: Lucia Sweeney
titled: The role of sleep in retention and generalization of statistically learned language

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Date: Aug 31, 2020
Rebecca L Gomez

Date: Sep 15, 2020

Date: Aug 31, 2020
Vicky Tzuyin Lai

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Date: Aug 31, 2020
Rebecca L Gomez
Dissertation Committee Chair
Psychology

Acknowledgements

With great pleasure, I would like to thank several people who supported me throughout this dissertation and my general studies at the University of Arizona.

First, this work would not have been possible without the guidance and support of my advisor, Dr. Rebecca Gomez. Thank you for inspiring me to pursue a graduate education, along with your invaluable advice and your encouragement during both my undergraduate and graduate career. Similarly, I could not have completed this dissertation without the aid of my dissertation committee as a whole, including Dr. Rebecca Gomez, Dr. Kenneth Forster, and Dr. Vicky Lai. Thank you for your incredibly helpful suggestions, valuable insight, and contributions to my research.

I would also like to thank my fellow graduate students (some now graduated) from the Child Cognition Lab. My research and overall experience in graduate school would not have been the same without you.

Dedication

I dedicate this dissertation to my loving family. Thank you for your unconditional emotional support, humor, and encouragement.

Table of Contents

List of Figures and Tables

List of Tables

Abstract

Across the lifespan, learners capitalize on regularities in language to find words in fluent speech and learn grammatical patterns, a process called statistical learning. Much of this research has focused on encoding, and few studies have investigated whether this type of knowledge is long-lasting. Related to retention, sleep promotes stabilization of memories in adulthood, with a growing literature on the benefits of sleep for memory in infants and children. In 5 studies, we examined the role of sleep in retention and generalization of statistically learned language. Specifically, 13-month-olds learned the order of syllables within words, discriminating words with the correct order from words with a reversed syllable order (Exp. 1), but they do not retain this level of detail across a nap (Exp. 2). Looking at the same type of learning in adults (Exp. 3), the group who slept after learning failed to show retention. Further analyses revealed an association between total sleep and memory for syllable order, indicating more sleep after learning related to retention. On the other hand, a group of wakefulness participants remembered which syllables formed words but not their correct order, and correlations showed they retained order after wakefulness if they encoded it during learning. Lastly, we demonstrated that sleep stabilized memory for grammatical patterns in 18-month-olds, promoting specific memory for learned phrases (Exp. 4). Additionally, only infants who napped after learning generalized their knowledge of the pattern to completely novel phrases (Exp. 5). Overall, our data suggest sleep does not benefit detailed representations earlier in infancy, but it promotes specific *and* abstract memory by 18 months. Furthermore, sleep relates to retention of order in adults, though wakefulness participants also showed retention. Future studies can elucidate how the effect of sleep develops, and whether sleeping after statistical learning in adulthood promotes memory for order more so than wakefulness.

Chapter 1: Overview

1.1 Language Acquisition and Statistical Learning

Learning a language is a daunting task. Learners are faced with acquiring a new system of sounds, an entire vocabulary, and the language's syntax, which governs how words combine to create sentences. How do learners, infants or adults, take the first steps to acquire a language when it is completely unfamiliar to them?

Of great benefit to the learner, language contains many regularities and cues to promote language acquisition. For example, infants and adults are able to track distributional statistics in speech. This information helps them acquire many aspects of their language, such as its phonetic inventory (Maye et al. 2002), lexical items and categories (Graf Estes et al., 2007; Hay et al., 2011; Maye et al., 2002; Saffran et al., 1996) and grammatical patterns (Gómez, 2002; Newport & Aslin, 2004; Thompson & Newport, 2007). This mechanism is referred to as statistical learning, our ability to extract statistical regularities and sequential dependencies existing in the input to form associations between related constituents.

For example, children monitor the co-ocurrence frequency between verbal labels and individual objects. They learn which labels often co-occur with specfic objects, allowing them to form associations between lexical items and their referents (Smith & Yu, 2008; Vlach & Johnson, 2013). Furthermore, infants and adults track co-occurrences between syllables to determine which combine to form words and which syllables span word boundaries (Saffran et al., 1996). Researchers refer to this phenomenon as word segmentation, and it is the focus of the first 3 Experiments in this report. Beyond ajdacent dependencies, infants also form associations between related items that do not occur immediately in sequence when the interleaving elements are highly variable (Gómez, 2002). We focus on acquisition of these nonadjacent dependencies in the final two Experiments of this report.

While many researchers have investigated the power of statistical learning in language acquisition, few have tested whether knowledge gained through this process is long-lasting. Once a language user has acquired new information through statistical learning, how long can they retain their knowledge? We address this gap in the literature by implementing a delay between exposure to a novel language and the test of participants' memory. Furthermore, an abundance of evidence suggests that sleep benefits memory retention in adulthood, though few studies have investigated whether sleep functions similarly early in development. To address this, we ask whether sleep supports retention of information acquired through statistical learning in both infancy and adulthood (Chapters 2-4). Specifically, we examine whether sleeping soon after learning promotes detailed representations of new knowledge gained through statistical learning, and whether the effect of sleep on retention changes with development (Chapters 2-6).

1.2 Sleep and Memory

A large body of literature supports the notion that there are sleep-dependent neural processes which stabilize and consolidate memories. In general, adults who sleep soon after learning will show better retention after a delay in comparison to individuals who stay awake for the same duration. This finding has been replicated with many types of learning, such as word learning (Dumay & Gaskell, 2007; Tamminen et al., 2010), spatial memory and navigation (Nguyen et al., 2013), and motor-based tasks (Doyon et al., 2009; Walker et al., 2002). What mechanisms might drive this effect? Research with adult participants revealed synchronous activity during sleep among subcortical structures, like the hippocampus, and cortical structures such as the neocortex (Born & Wilhelm, 2012). The dialogue between these regions likely supports the transfer of information from short-term memory to long-term memory, promoting retention (Born & Wilhelm, 2012; Diekelmann & Born, 2010).

More specifically, results from animal studies using single-cell recordings suggest that one of the mechanisms driving sleep-dependent memory consolidation occurs in the hippocampus, a subcortical structure located in the temporal lobe. During learning, neurons fire in specific sequences that correspond with the information being acquired. During sleep, the same neurons fire in the patterns observed during learning (Pavlides & Winson, 1989; Skaggs & McNaughton, 1996). Researchers refer to this phenomenon as *neural replay*, when patterns of activation that occurred during learning are replayed in the brain during subsequent sleep. Across slow-wave sleep, these bursts of activity in the hippocampus occur in synch with sleep-spindles, generated in the thalamus. In turn, sleep-spindles occur within the down-phases of slow-oscillations, which are generated throughout the cortex. This synchronous activity assists the transfer of information from the hippocampus into cortical memory (reviewed in Born & Wilhelm, 2012).

While sleep promotes memory consolidation in adulthood, fewer researchers have examined how the effect of sleep develops as the brain and its systems mature. Early in development, young infants have fewer established neural connections. The networks that are in place at this stage may not support the processes behind sleep-dependent memory consolidation to the same extent as in older children, and adults. In other words, the role of sleep in learning and memory may change as the infant brain develops (Gómez & Edgin, 2015). Following this line of thought, it is important to question whether sleep functions the same way across development. In infancy, our brains are still developing while we are tasked with the difficult challenge of learning everything about the world, including acquiring our primary language.

Young infants also take several naps throughout the day, only waking for a few hours at a time. Could these naps function to help young learners remember what they encountered during their periods of wakefulness?

1.2.1 Sleep and Non-Linguistic Memory in Infancy

In fact, a number of studies demonstrate the beneficial effect of sleep on memory early in life. Specifically during maturation when the brain is highly plastic, sleep may assist neural reorganization after periods of learning. Frank et al. (2001) discovered that when kittens slept after a period of monocular deprivation, they showed twice the amount of neural changes than kittens that were kept awake after an identical deprivation experience. Additional studies revealed cellular processes occurring during sleep, including reactivation of the cortical regions active during learning (Jha et al., 2005), promote strengthening of synaptic connections formed during wakefulness (Aton et al., 2009; Seibt et al., 2012).

Furthermore, behavioral research with human infants indicates sleep confers a benefit to memory retention in children younger than 2 years of age, a time of great development. A recent study examining sleep's effect on memory in infancy used a deferred imitation learning task, which involves nonverbal explicit learning (Seehagen et al., 2015). The authors performed a sequence of actions using a puppet while an infant participant, aged either 6 or 12 months, observed. Following the researcher's demonstration, infants either completed the test immediately, napped after learning, or stayed awake after learning. During the test, the researcher gave the infant the same puppet from the demonstration, and the infant had 90 seconds to spontaneously imitate the actions they had observed earlier. The authors found a boost in memory with sleep, with infants who napped performing significantly more actions that the group tested immediately. In fact, infants who napped and were tested 24 hours later still recalled more actions than infants tested immediately after learning. The authors argue that napping after learning stabilizes the recently encoded memories, so that they are protected from interference until they are more fully consolidated during nighttime sleep (Seehagen et al., 2015). This study provides substantial evidence that sleep benefits memory for a sequence of actions in very early childhood, as early as 6 months of age.

With more simple learning materials, researchers have observed a benefit of sleep on memory in infants as young as 3 months of age. After viewing a novel cartoon face, 3-month-old infants can only remember the face 2 hours later if they nap during the delay. Furthermore, the density of sleep spindles during their post-learning nap associated with their habituation time. The authors reason that habituation time reflects general cognitive abilities, so sleep spindles may be an indication of or precursor to early intelligence (Horváth et al., 2018). Together, these studies provide evidence that sleep promotes memory retention early in life.

Sleep additionally benefits retention of motor-based problem solving in infancy. Berger and Scher (2017) used a task that required walking infants (9-16 months of age) to crawl through a tunnel. This is challenging because once infants are walking, they are unlikely to revert to crawling. The authors found that infants who slept after attempting to solve the problem showed an increase in performance at the delayed test. The infants who napped crawled through the tunnel faster, needed less instruction, and made fewer errors than they did before their nap. The wakefulness group, on the other hand, showed no change in performance across the delay containing only wakefulness. It appears that sleep had an influence on problem solving ability, allowing infants to integrate across their earlier learning experience and realize a solution more quickly than infants who did not sleep (Berger & Scher, 2017).

Overall, it is clear that sleep provides a benefit for memory retention, even early in life. However, these studies have assessed problem solving and retention of declarative memories, which may involve different learning systems than language acquisition and statistical learning. In the following section, we review existing research supporting the idea that beneficial effects of sleep extend to language acquisition throughout the lifespan.

1.2.2 Sleep and Memory for Language

1.2.2.1 Word Learning. Relevant to our questions regarding retention after statistical language learning, researchers have specifically examined whether sleep benefits memory for language in both childhood and adulthood. Many of these studies involve word learning tasks measuring encoding, consolidation and integration of new vocabulary into the existing lexicon. Relevant to retention in adults, Gaskell and colleagues use lexical competition tasks to assess whether new words are integrated into learners' vocabularies. For example, if a participant learns the novel word "banara," once this item has become part of the lexicon, it will compete for activation with "banana" when the participant hears the syllable "ban-." Such competition leads to longer reaction times in processing the word "banana."

Dumay and Gaskell (2007) discovered a specific role of sleep in the process of lexical integration. They found that newly learned words engage in lexical competition with familiar words only once participants sleep after learning; participants who stay awake after learning do not show evidence of lexical competition. Sleep also benefitted participants' ability to recall the novel words freely, without prompting (Dumay & Gaskell, 2007). This result also extends to visual word recognition, when novel words are presented orthographically and not auditorily (Wang et al., 2017), though in both visual and audial domains sleep preferentially integrates novel words that were strongly encoded over word forms that were weakly encoded (Walker et al., 2019). Lastly, researchers found that the number of sleep spindles generated during sleep positively correlates with the lexical competition effect (Tamminen et al., 2010). Combined, these findings suggests sleep plays an active role in the integration of new vocabulary into the existing lexicon (Tamminen et al., 2010).

Does sleep have a similar role in children's word learning? Henderson and colleagues replicated Gaskell's adult findings, revealing that sleep promotes lexical integration in children ranging from 7-12 years of age (Henderson et al., 2012). Furthermore, outside of lexical integration, Williams & Horst (2014) assessed whether 3-year-old children learn novel words via storybooks, and whether sleep promotes retention of the new vocabulary. Overall, the authors found a beneficial effect of sleep: the children who napped showed greater retention of the words than children who stayed awake. This result persisted 24 hours and 7 days after children read the storybook. Such long-lasting effects indicate sleeping soon after learning facilitated the consolidation of the words to long-term memory (Williams & Horst, 2014).

Even in infancy, napping after learning new object-label mappings helped 16-month-olds retain those mappings over a delay averaging 2 hours (Horváth et al., 2015). Infants who napped showed a marked *improvement* in their fixations to the target object, while infants who stayed awake did not demonstrate a change in accuracy. Furthermore, within the nap group, infants who had a relatively large vocabulary as measured by the Oxford Communicative Development Inventory benefitted the most from sleeping after learning (Horváth et al., 2015). This pattern of results suggests the amount of pre-existing knowledge a child has influences the degree to which they benefit from sleep. As infants develop and their knowledge expands, we may see a larger benefit from sleep. In addition to behavioral studies, research using ERPs shows sleep promotes retention and generalization of newly learned words in infancy Friedrich and colleagues (2015) found that after napping, 9- to 16-month-old infants retain the recently learned labels and apply them to novel exemplars. The infants who stayed awake in the hours after learning did not retain the words, and they could not recognize new exemplars of the learned category. It appears that even in infancy, sleeping after a learning experience encourages generalization so that learners can apply their knowledge to novel instances of the learned information.

1.2.2.2 Sequences and Grammar. Apart from word learning, sleep also plays an important role in the retention and generalization of grammatical rules acquired through statistical learning. Specifically, in adulthood, targeted memory reactivation during sleep facilitates grammatical abstraction (Batterink & Paller, 2017). In this study, adults listened to an artificial language containing 20 monosyllabic words presented in 56 different phrases. The words were split into six categories, and each lexical category could only occur in specific positions within each phrase. Participants slept after learning, and the researchers presented the artificial language to half of the participants during slow-wave sleep. This process is thought to promote neural replay of learned material during sleep (Rasch et al., 2007). The other half were not cued during sleep. The participants who were cued with the grammar showed increased generalization at test, recognizing the correct order of words in new phrases that were not presented during learning. This study provides evidence that replay of recently learned memories during sleep promotes learning and abstraction of grammar (Batterink & Paller, 2017).

In addition, sleep plays an important role in generalization and rule abstraction, with slow-wave sleep (SWS) and rapid eye movement sleep (REM) functioning together to promote grammatical abstraction (Batterink et al., 2014). The authors presented an artificial language to adult learners. In this study, adults read two-word phrases. The researchers informed participants that the first word in each phrase is an article (*gi, ul, ro,* and *ne*) and the second word was a familiar noun, half were animate and half inanimate. The researchers also specified that two of the articles indicate the noun is nearby (*gi* and *ro*), while the other two articles indicate that the noun is far away (*ul* and *ne*). They did not inform the participants that two articles were consistently paired with animate nouns (*gi* and *ul*) and the other two paired with inanimate nouns (*ro* and *ne*).

On each trial of the learning phase, participants read a single phrase and indicated (1) whether the noun was relatively near or far and (2) whether it was animate or inanimate. They were provided feedback after each response. Participants who slept after the initial learning phase showed greater rule abstraction after the delay, signified by greater response times on "switch" trials, in which an article was *incorrectly* paired with a noun according to the animacy rule (i.e. an article marking animate nouns paired with an inanimate object). Furthermore, participants who spent more equal amounts of time in SWS and REM sleep showed the strongest evidence of grammatical abstraction. This pattern of results indicates that sleep benefitted their representation of this rule at an abstract level (Batterink et al., 2014).

Moreover, sleep benefits grammatical generalization in infants and children. The work reviewed here focuses on statistical learning of nonadjacent dependencies (NADs) in language. For example, in English we have phrases like, "he is jumping," and, "she is playing," where "is" and "-ing" remain consistent but the rest of the vocabulary varies. In these phrases, "is" and "-ing" form an NAD. Using ERP, Schaadt and colleagues (2020) revealed a change in how children represent NADs across a period of sleep. Immediately after listening to a language containing consistent NADs, children performed a grammaticality judgment task while the researchers measured ERPs. During the test, ungrammatical phrases elicited a positive response occurring 800-1000ms after stimulus offset. Contrastingly, after a period of 24 hours including sleep, ungrammatical phrases elicited a negative ERP response 600-1000ms after stimulus offset. Grammatical violations typically elicit negative ERP responses in adults, but when their prefrontal cortex is inhibited (via transcranial direct-current stimulation) they tend to show a positive response (Friederici et al., 2013). Considering these effects found in adults, the authors argue that the negative response observed on day 2 is a sign of more consolidated processing, potentially reflecting representation of NADs in lexical memory after sleep (Schaadt et al., 2020).

Lastly, infants also observe a benefit from sleep when it comes to abstraction of grammatical patterns. Gomez and colleagues (2006) first familiarized 15-month-old infants with an artificial language containing two NADs. This language consisted of three-word phrases in which the first word was consistently paired with the third word. Infants either slept or remained awake after learning, following their typical sleep schedule. Four hours after familiarization, infants completed a test in head turn preference. The infants who stayed awake showed veridical retention, discriminating trials violating the NADs from familiarization from the trials containing NADs consistent with familiarization. Alternatively, infants who napped generalized the NAD pattern: they remembered the predictive relationship between the first word and third word in each phrase (Gómez et al., 2006). Furthermore, a follow-up study demonstrated that infants who do not sleep after learning do not retain their veridical representations across 24 hours. On the other hand, infants who sleep continue to show generalization effects, indicating sleep promoted a long-lasting abstract representation of the NADs in 15-month-olds (Hupbach et al., 2009).

1.3 Overview of Present Studies

The existing research on sleep, memory and language clearly demonstrates the benefits of sleep on subsequent retention and even abstraction of learned information, especially in regard to word learning and generalization of grammatical rules. However, few have examined the roll of sleep in retention of words learned via word segmentation. Moreover, there are few studies assessing whether the effects of sleep change with neural development and maturation.

We are contributing to this field by continuing the line of research on retention of statistically learned language. In 3 studies, we use a word segmentation task to assess whether infants and adults retain detailed representations of the learned words. More specifically, we measure whether learners (adults and infants) encode *and* retain the correct order of syllables within words. Then, in 2 studies, we examine the role of sleep on veridical retention and generalization of nonadjacent dependencies in 18-month-old infants. As Gomez and colleagues conducted a similar study looking at acquisition of nonadjacent dependencies in 15-month-olds, we compare the pattern of effects of sleep observed in Gomez et al (2006) to our present results with 18-month old infants. Across these 5 studies, we hope to elucidate the role of sleep in consolidation of statistically learned language and determine whether sleep provides the same benefits to both younger and older learners.

Chapter 2: Experiment 1

Acquisition of Temporal Order During Word Segmentation in Infancy

2.1 Introduction

When learners encounter a language that is completely new to them, it is incredibly difficult to locate single words in natural, fluid speech. Rather, the words run together and are not clearly separated from one another because speakers do not pause between each word when talking naturally. How might a learner go about segmenting a novel speech stream into single words?

In a seminal study, Saffran, Aslin, and Newport (1996) demonstrated that young infants are able to complete this task, referred to as word segmentation, using sequential dependencies within the speech stream. They created an artificial language that contained four three-syllable words: *bidaku, padoti, golabu,* and *tupiro.* The language contained no indication to the location of word boundaries, such as pauses or changes in lexical stress, so that the only cue to the location of word boundaries was the conditional probabilities between syllables. Conditional probabilities reflect the likelihood that one syllable will follow another and are calculated by taking the frequency of syllable pair XY and dividing it by the frequency of syllable X (Saffran et al., 1996). In Saffran's language, syllables that combined into words had conditional probabilities equal to 1. To illustrate, syllable *bi* was always followed by *da,* so these two items had a conditional probability of 1. *Da* was always followed by *ku,* so these syllables also had a conditional probability of 1. *Ku,* the final syllable in *bidaku*, could be followed by one of three syllables, either *pa, go,* or *tu*, as there was an equal likelihood that *padoti, golabu,* and *tupiro* would be the next words in the sequence. Therefore, the conditional probability between *ku* and *pa, go,* and *tu* were each .33.

Both young and mature learners use this information to determine which syllables combine to form words and which syllables span word boundaries. After exposure to a language, learners are able to distinguish syllable sequences with relatively low conditional probabilities (part-words, like *kupado* from the language above) from syllable sequences that contain high probabilities (words). In Saffran's original language, all word tokens appeared at an equal frequency. This resulted in words being more frequent than part-words in the corpus, so at test learners could have discriminated based on the co-occurrence frequency of syllables (how often a sequence of two syllables occurs) rather than conditional probabilities. Aslin, Saffran, and Newport (1998) followed up by equating the frequency of words and part-words in their familiarization language. Even when part-words are equally as frequent as words in the corpus, infants distinguish between them. These results establish that language learners rely on conditional probabilities over the mere co-occurrence frequency of syllables (Aslin et al., 1998). In our studies examining word segmentation, we are interested in learners' retention of syllable order and not conditional probability per se. Thus, we did not control for co-occurrence frequency, and all of our words appear equally often in our languages.

Further research on statistical word segmentation indicates that both infants and adults represent syllable sequences with high-conditional probabilities as potential word forms. Young and old learners are more likely to map high-probability strings as labels for novel objects compared to low-probability strings (Graf Estes et al., 2007; Mirman et al., 2008). Furthermore, statistical learning applies to acquisition of natural languages, as English-speaking infants can use conditional probabilities in order to segment words in completely novel Italian utterances (Hay et al., 2011).

In the present study, we ask whether 13-month-old infants encode temporal order of syllables during statistical word segmentation. Specifically, we measure whether they can discriminate words with the correct syllable order from words with an incorrect syllable order immediately after listening to an artificial language. While the main focus of this report is the effect of sleep on retention of statistical learning, we first conducted Experiment 1 to establish that infants encode syllable order during word segmentation.

To date, few researchers have empirically investigated whether infants track temporal order during aural statistical learning, though the notion of conditional probability includes an assumption of order (syllable X *is followed by* syllable Y). To expand this vein of research, Pelucci and colleagues (2009) assessed whether infants track backwards conditional probabilities in natural speech. In this study, 8.5-month-old infants listened to sentences spoken in Italian. The corpus contained the target words *fuga, melo, casa,* and *bici.* Critically, while these words all had forwards conditional probabilities of 1.0, syllables *sa* and *ci* both occurred elsewhere the language, in addition to within the words listed. These extra instances caused the backwards conditional probability of *casa* and *bici* to drop to 0.33, as *sa* and *ci* each had two other syllables come immediately before them in the corpus. The syllables in *fuga* and *melo*, however, did not occur elsewhere in the language: *ga* was always preceded by *fu* and *lo* was always preceded by *me*. Therefore, these items had a backwards transitional probability of 1. At test, the researchers assessed whether infants could discriminate words with high backwards transitional probabilities (*fuga* and *melo*) from those with low backwards transitional probabilities (*casa* and *bici*). Infants discriminated the word forms based on their backwards conditional probabilities, listening longer to words with high backwards conditional probabilities than low backwards transitional probabilities. This study is a clear demonstration that infants monitor regularities in the backwards direction in addition to the forwards direction (Pelucchi et al., 2009).

Furthermore, infants track ordinal information in the visual domain. For example, after exposure to visual sequences following patterns such as: ABAA, CBCC, and DBDD, 10-month-old infants learn item B fills the second position in a sequence of this form. They can also recognize similar patterns even when the surrounding repeated constituent is novel (Lewkowicz, 2013). Additionally, de Hevia and Spelke (2009) familiarized infants with either an ascending or descending sequence of displays. In the ascending displays, the screen began with a small number of dots, which increased in number throughout the sequence. The descending display began with a large number of dots, which decreased in number throughout the sequence. After habituating to one of the sequences, infants look longer at test stimuli with the reversed ordering. If infants habituated to an ascending series, they attended longer towards a descending series at test, and vice versa. This indicates that infants are able to determine a familiar, correctly ordered sequence from the same sequence in the opposite order (de Hevia & Spelke, 2010).

We plan to manipulate temporal order at test to assess the level of detail in the lexical representations formed during statistical learning. Namely, we will test whether infant learners encode the order of syllables within words during word segmentation (Experiment 1). In Experiment 2, we follow up our results from Experiment 1 and assess whether infants retain syllable order if they nap soon after learning. Experiment 3 focuses on adult learners, questioning whether participants who sleep after learning will show greater retention of order than those who stay awake. Experiment 3 also assesses the role of immediate testing in long-term retention.

2.2 Methods

We recruited 24 13-month-old infants from monolingual-English households to participate in this study. They had an average age of 13.1 months (range 12.23-13.99 months). The researchers scheduled all infants to participate at a time they were expected to be awake and well-rested. We did not include infants in our final sample if: parents or siblings interfered during familiarization or test procedures (4) there was experimenter error (2), or infants became fussy and could not complete at least half of the test session or 6 trials (1).

2.2.1 Materials

Familiarization Stimuli. We created three pairs of languages (Set A, Set B, and Set C). We wanted to ensure maximum generalizability, so we created a total of six languages. All languages contained four disyllabic words that were phonotactically correct in English. Within each set, the two languages were counterbalanced so that legal items in Language 1 (L1) were illegal items in Language 2 (L2) and vice versa. For example, Set A L1 contains the words *diro* and *lapi,* while L2 contains *rodi* and *pila.* The order of their syllables has been reversed between languages, and legal items in one language are backwards-order (illegal) items in the other. Furthermore, L1 has the words *budo,* and *tani* while L2 has *dota* and *nibu*. We created these words in L2 by taking the words from L1 and combining the last syllable of one word with the first syllable of the other. This generates items with relatively low sequential dependencies, which we call part-words. For these items, the legal words in one language are the part-words in the other. The vocabulary in each language is displayed in Table 1. It is important to note that we do not test infants' discrimination of part-words from words in Experiments 1 and 2. Rather, we focus only on infants' discrimination based on syllable order, distinguishing forwards words from backwards-order items. We will assess discrimination of the part-words in future studies with infants, and we use these items when testing adult subjects in Experiment 3.

Table 1.

Set A		Set B		Set C	
L1	L2	L1	L2	L1	L2
diro	rodi	bilo	lobi	gobi	bigo
lapi	pila	daki	kida	dira	radi
budo	dota	doro	romi	sila	nusi
tani	nibu	miga	gado	monu	lamo

Table 1 displays the three language sets used in Experiments 1 and 2. Each set was constructed so that L1and L2 were balanced, with 2 words from L1 forming backwards-order words in L2, and vice versa. We included this manipulation to ensure infants were not responding based on preferences for specific word forms.

We created all of the familiarization languages so that there were no pauses or changes in stress or pitch throughout the speech streams. Therefore, there were no indications to the locations of word boundaries aside from the sequential dependencies between syllables. To ensure the participants could not use the beginning and end of the stream as a segmentation cue, the audio files were set to begin silently, slowly ramp up in volume, and slowly reduce in volume at the end. All languages lasted approximately 13 minutes. It is important to note that Saffran and colleagues' (1996) familiarization language lasted 2 minutes; we extended familiarization in these studies to ensure infants had ample time for encoding. Researchers have not yet established how the length of exposure influences encoding and subsequent retention, and our decision to use a 13-minute familiarization is not necessarily justified by previous literature (though this a potential direction for future research).

Test Stimuli. At test, infants completed 12 trials using the Head-Turn Preference Procedure (Kemler Nelson et al., 1995). Half of the trials contained familiar stimuli and presented forwards-order words (FW) from the familiarization language – such as *budo* and *tani* from L1. The other half of the trials were unfamiliar and contained backwards-order words (BW) – such as *rodi* and *lapi,* which were created by reversing the order of the syllables in L1 words *diro* and *pila*. On each trial, the infant heard a single item repeated until they displayed a lack of interest (procedure described in more detail below) or until the trial ended automatically after 20 seconds. We compared looking time during trials containing FWs to looking time during trials containing BWs.

2.2.2 Procedure

During familiarization, the infants played quietly in the lab space with the researcher and the caregiver. The researcher instructed the caregiver not to speak while they were listening to the language unless their child became uncomfortable and they would like to stop the experiment. Immediately after familiarization, the researcher escorted the infant and caregiver into the Head Turn Preference booth. The caregiver wore noise-canceling headphones, and the researcher instructed them to keep their child sitting facing forwards on their lap and to avoid interacting with their child in any capacity unless they became fussy.

In the Head-Turn Preference Procedure, infants were seated facing forward on their caregiver's lap in a sound-attenuated booth. The researcher instructed the caregiver to keep their child sitting facing forwards and to refrain from interacting with them, directing their attention, or speaking during the test session. Once the infant and caregiver were situated, the researcher sat in an adjacent room to observe the participant over closed-circuit television and control the experimental program. In the participant testing booth, there was a light fixed in the center of the wall immediately in front of the participant. There were also single lights fixed on the walls to the left and right of the participant. Both of these side lights had a speaker attached to the wall directly below the light. At the start of each trial, the center light began blinking. Once the infant fixated on the center light, the researcher pressed a button which extinguished the center light and illuminated one of the lights to the side of the participant. The side light blinked until the infant turned their head 30 degrees towards that direction. Once the infant fixated on this light, the researcher pressed another button, initiating the presentation of audio stimuli through the speaker adjacent to the blinking light. The grammar played through the speaker until the child looked away from the light for two consecutive seconds or until the audio file ended, whichever occurred first.

2.3 Results

Our main measure if interest is whether infants looked longer to one trial type (BW or FW) than the other. This would indicate they encoded syllable order and discriminated the trials containing forwards words from the trials containing backwards-order words. To begin analysis, we first calculated the average total looking time to FW trials and BW trials separately. Then, we generated a difference score for each infant by taking their average looking time to FW trials and subtracting their average looking time to BW trials. A value significantly varying from zero indicates infants were able to discriminate the legal FW items from the BW items, which violated the sequential dependencies heard during familiarization.

In order to ensure the language set (A, B, or C) did not influence learning, we assessed whether infants' performance at test differed based on what language they listened to during familiarization. We conducted a two-way ANOVA with Language Set (A, B, or C) and Gender as our factors and the infants' difference scores as our dependent measure. There were no significant main effects or interactions [$F(2,18) = 0.681$, $p > 0.05$], so we collapsed across

Language Set and Gender for the following analyses. See mean looking times to FW and BW trials in Figure 1.

Then, we conducted a paired-sample t-test, comparing infants' looking time to FW trials to their looking time to BW trials. We found a significant difference in looking to the two trial types [$t(23) = 2.674$, $p = 0.014$, $d = 0.546$], with longer average looking to FW trials ($M = 8.61$s) than BW trials ($M = 7.19$s). The infants discriminated the words with the correct order (in FW trials) from the words with the incorrect syllable order (in BW trials).

Figure 1.

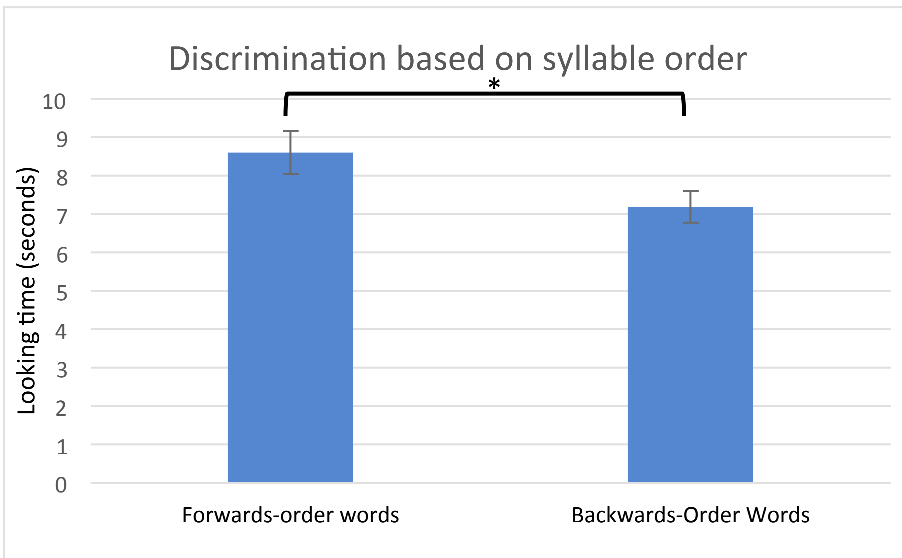

Figure 1 displays infants' average looking time to trials containing forwards-order words and trials containing backwards-order words. Error bars reflect standard error of the mean (SEM). We found a significant difference, with infants looking significantly longer to trials presenting words.

2.4 Discussion

We hypothesized that infants would encode syllable order during word segmentation. Our results are in support of this hypothesis. We found infants looked longer during trials containing forwards-order words than they did during backwards-order trials. In other words, they remembered the words heard during familiarization (along with the order of syllables within them) and discriminated the forwards items from the backwards items at test.

While infants in this study were successful in discriminating the correct from incorrect syllable order, they demonstrated a familiarity effect, listening longer to trials that were consistent with the language they listened to during familiarization. Contrastingly, Saffran et al. (1996) reported a novelty effect. Infants in their study listened longer to trials containing part-words than trials containing words. This is in accordance with the general reasoning behind habituation effects in infant research: as infants learn about a stimulus, they pay less attention to it as time goes on. Once infants show a significant decrease in looking, if the stimulus changes to something novel the infant will show an increase in attention to the novelty. Therefore, researchers have suggested novelty preferences indicate complete encoding of the familiarization materials, while a familiarity preference indicates the infant was still in the process of learning about the familiarization stimulus (Houston□Price & Nakai, 2004). This is supported by research showing relatively short exposures to a stimulus will produce a familiarity effect, while long exposure times correspond to novelty effects (Hunter et al., 1983; Roder et al., 2000). Furthermore, after learning about a visual stimulus, infants will show novelty preferences soon after learning but familiarity preferences after a delay of 1 month (Courage & Howe, 2001). Our familiarity effect could indicate that the infants had not strongly encoded the order of syllables within words. However, they learned enough to recognize the test trials that were consistent with their familiarization language, and they attended longer to these trials than the backwards-order trials. We continue this line of research by questioning whether 13-month-old infants retain syllable order across a delay containing sleep.

Chapter 3: Experiment 2

Does Sleep Benefit Retention of Temporal Order in Infancy?

3.1 Introduction

It has been widely established that learners use sequential information, such as distributional regularities, to locate words in a stream of speech as discussed in Chapters 1 and 2. However, only a small number of studies focus on long-term retention of these word forms, which is necessary for learners to benefit from their ability to track conditional probabilities. To expand on Experiment 1, we conducted a study using the same materials but added a delay including sleep between the time of exposure to a novel language and test. This will allow us to determine whether infants can retain temporal order across a period of sleep.

In adult research, a widely replicated finding demonstrates that adults who sleep after learning new information will show greater long-term memory retention than adults who stay awake after learning (reviewed in Stickgold, 2005). Relevant to development, studies with kittens suggests sleep has a role in promoting neural plasticity after a cortical learning experience (Aton et al., 2009; M. G. Frank et al., 2001; Seibt et al., 2012), and behavioral studies indicate sleep promotes retention of memories for faces as early as 3 months of age (Horváth et al., 2015). To date, the only study focusing on retention of segmented words in infancy suggested sleep benefits infants as young as 6.5 months (Simon et al., 2017). However, infants only discriminated words with high sequential dependencies from low sequential dependencies in the first block of testing. As infants heard part-words repeated throughout the test, these items may have interfered with their memories for the words. This suggests while sleep benefits memory for statistical information early in life, the retained representations are still fragile and prone to interference. Additionally, infants with more mature sleep physiology, as measured by polysomnography, demonstrated more robust retention at test. This relationship provides evidence that sleep has a role in memory consolidation early in life, especially as infants mature (Simon et al., 2017).

In the present study, we ask whether 13-month-old infants retain the correct order of syllables within words across a four-hour delay. Adult studies focusing on retention of order indicate sleep benefits sequence-based motor skill learning (Walker et al., 2002). Sleep also appears to preferentially strengthen forwards transitions rather than backwards transitions. For instance, Drosopoulos and colleagues taught participants word triplets (e.g. chess, winter, and traffic). Within each triplet, the words were always presented in the same order. After a period of sleep, participants showed enhanced memory for forwards transitions only. In other words, when presented with the first word in a triplet at the delayed test, they recalled the second word more often than the wakefulness group despite equal levels of encoding before the delay. Sleeping promoted consolidation of forwards transitions, though sleep and wakefulness groups were equal in their ability to recall backwards transitions (generating the second word when provided with the last word in a triplet). Furthermore, the sleep group recalled forwards transitions more often than backwards transitions, suggesting sleep preferentially facilitates consolidation of forwards transitions but not backwards transitions (Drosopoulos et al., 2007). This study focused on declarative learning and sequencing of events, though a similar process may be active during or after statistical learning. If sleep benefits memory for order, it could be the case that once infants segment a word from the speech stream, like “rodi,” sleep might strengthen the transition between “ro” and “di,” but not the backwards transition between “di” and “ro.” Therefore, we have reason to believe infants (and adults who sleep in Experiment 3) may show retention of syllable order across sleep.

3.2 Methods

Data collection for this study has halted temporarily because of the coronavirus pandemic. At this point in time, we have recruited 8 13-month-old infants (*M* = 13.23 months, range = 12.62 – 14.1 months) from monolingual-English households. We plan to include a total of 24 infants in our sample for this study. We scheduled all infants so that they were expected to nap within the four-hour delay between learning and test.

3.2.1 Materials

Familiarization Stimuli. The familiarization languages in this study were identical to those used in Experiment 1. Each infant listened to one of six languages. Each language contained four syllabic words and included no auditory cues to the location of word boundaries.

Test Stimuli. The test stimuli in this study were identical to Experiment 1. The test contained 12 trials, half presenting words from the familiarization language and half presenting backwards-order words.

3.2.2 Procedure

Familiarization. The familiarization phase took place in participants' homes. An assistant accompanied the researcher to play with siblings in a separate room. Before beginning familiarization, the researcher and caregiver found a quiet place to complete the study. The researcher also instructed the caregiver not to speak while the language was playing unless their child became fussy and they would like to stop the study. The researcher and caregiver played with the infant nonverbally with quiet toys while listening to the language.

Test. Four hours after familiarization, the participating families came to the lab to complete the test. As in Experiment 1, we used The Head Turn Preference Procedure (Kemler Nelson et al., 1995) to assess infants' memory for the order of syllables within words.

3.3 Results

First, we ruled out effects of gender and Language Set. We calculated difference scores by subtracting the average looking to BW trials from the average looking to FW trials. A positive difference score indicated infants looked longer to FW trials. We then conducted a univariate ANOVA with Language Set (A, B, or C) and Gender as between-subjects factors. We found no main effects or interactions [$F(1,3) = 2.84$, $p > 0.05$], so we collapsed across these variables for the remaining analysis. See looking times to FW and BW stimulus types in Figure 2.

Like in Experiment 1, our main measure of interest was whether infants looked longer to one trial type (FW or BW trials) versus in the other. We conducted a two-tailed, paired-samples t-test comparing listening time to FW trials against listening time to BW trials. There was no significant difference in their listening time ($M_{FW} = 7.12$s, $M_{BW} = 7.56$s; $t(7) = -0.407$, $p > 0.05$), indicating they did not discriminate words with the correct syllable order from words with the incorrect syllable order.

Figure 2.

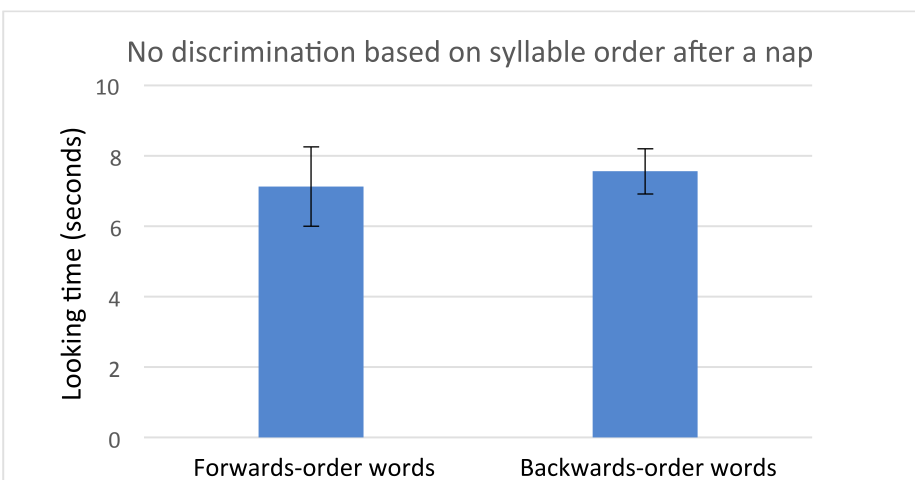

Figure 2 illustrates infants' listening time to words and backwards-order words. Error bars represent SEM. We found no significant difference in their listening times to these types of trials, indicating they could not discriminate the backwards words from the forwards words. This finding suggests they did not remember the correct order of syllables within words.

We followed up by comparing the results of Experiment 1 to Experiment 2, conducting a repeated-measures ANOVA with Group (Immediate or Nap) as a between subjects factor and Trial Type (Words and Backwards-order words) as a within-subjects factor. We found no main effects or interactions [$F(1, 30) = 2.82$, $p > 0.05$]. While there are currently no differences between performance in our immediate and nap conditions, we expect this test to be more reliable once we have collected more data in our nap condition.

3.4 Discussion

In our preliminary set of data, infants who nap after a word segmentation task do not retain syllable order. However, we found no significant differences between the behavior of infants in the nap group (from this Experiment) and the group of infants tested immediately in Experiment 1. We currently have a small sample size in the nap condition, so we will assess the difference between groups again once we collect more data. In addition, we do not have a group of infants who remain awake during the delay between learning and test, so we cannot compare retention between napping and wakefulness. Though considering Gomez and colleagues (2006) found 15-month-olds retain detailed representations of nonadjacent dependencies only across wakefulness (and not sleep), it's possible 13-month-olds may also retain knowledge of syllable order across wakefulness. As young infants encode word forms with a great amount of detail (i.e. including talker voice as part of their representation [Houston & Jusczyk, 2000]) it could be the case that sleep promotes loss of detail early in life.

To some extent, less-specific memories for language allow learners to extend their knowledge to novel sentences or speakers. Yet, language is inherently sequence-dependent: the correct order of sounds, syllables, words, and sentences is necessary for successful communication. Failure to retain the proper order of syllables within words doesn't advantage the learner in any way, though our sample size is still small. Continuing data collection in this condition will help clarify the role of sleep in the retention of syllable order in this age group.

In the future, we plan to follow up on these results by assessing whether infants can retain syllable order across a delay containing wakefulness. If our final group of nap infants continues to show lack of discrimination between FW and BW items, and a wakefulness group successfully retains syllable order, this pattern of results suggests sleep promotes forgetting of order-related information within words. We also plan to assess whether infants discriminate forwards words from part-words after delays containing sleep versus wakefulness. It could be the case that sleep promotes retention of co-occurrence frequency but not order in infancy. Furthermore, we plan to assess retention of syllable order in slightly older children, as it's possible that sleep will promote stabilization of details later in life. Infants who are further along in development have more mature neural systems and more pre-existing knowledge regarding language. Older infants' neural connections could support sleep-dependent consolidation to a greater extent than in younger infants, and previous studies indicate a beneficial role of pre-existing knowledge in sleep-dependent consolidation (Horvath et al., 2015). For these reasons, we might expect to see retention of order in older children across sleep.

Chapter 4: Experiment 3

Retention and Specificity of Lexical Representations After Statistical Learning in Adults

4.1 Introduction

As reviewed, much research has focused on infants' ability to use statistical information within language to segment strings of speech into single words. Investigators have also extended this line of questioning to language learning in adulthood, asking whether adult learners use sequential dependencies to segment words from fluent speech. Notably, Saffran, Newport, and Aslin (1996) presented undergraduate students with an artificial language. The language contained six trisyllabic words repeated in a pseudorandom order with no consecutive repetitions. Researchers synthesized the language so that there were no changes in pitch or stress, and there were no pauses between words. Thus, the only cue to the location of word boundaries was the frequency with which syllables occurred in sequence, or their co-occurrence frequency.

After listening to this language for a total of 21 minutes, participants completed a two-alternative forced-choice task. Each trial, they heard a word from the language along with an illegal item. There were two types of illegal items: part-words and nonwords. Researchers created part-words by taking a word and altering either the first or last syllable (e.g. the word "pidabu" was altered to form the part-word "pidata"). They also created nonwords by combining three syllables that never occurred in sequence within the language. On every test trial, participants needed to choose which of the two options sounded more like a word from the language they heard earlier. On average, they correctly selected the words significantly more than chance would dictate, showing they successfully extracted the word forms from the artificial language purely based on the sequential dependencies between syllables (Saffran et al., 1996).

Few studies have examined whether adults retain information gained through statistical learning. Frank and colleagues (2013) had adults listen to a large corpus of artificial speech containing 1000 different words. The corpus was presented in sentences, with 200ms pauses between sentences. Within sentences, there were no pauses between words and no prosodic indications to the locations of word boundaries. The participants listened to the language for one hour each day for 10 days. Three years after this 10-day exposure, participants scored with over 50% accuracy on a two-alternative forced-choice task pitting words against part-words, indicating they retained at least some of the words from the corpus (Frank et al., 2013).

It is important to emphasize that this study used a large-scale language and multiple, hour-long exposures, mimicking real-life experiences with natural language. Their findings strongly suggest that word-form representations gained through statistical learning can be long-lasting given sufficient exposure. Moreover, the authors did not track whether participants slept soon after listening to the language each day. While not the focus of Frank et al. (2013), it's possible that sleeping soon after exposure to a language could bolster retention further than the authors reported. As reviewed, a large body of literature demonstrates that adults who sleep after learning new information will show greater long-term memory retention than adults who stay awake (reviewed in Stickgold, 2005), and it is possible that the benefits of sleep extend to retention of statistically learned language.

As in Experiments 1 and 2 in this report, we are specifically interested in whether sleep benefits memory for temporal order after statistical learning. With respect to the role of sleep in retention of temporal order, the majority of studies have focused on order in visual or motor-based sequences. For example, sleep benefits retention of a finger-tapping sequence in adults (Doyon et al., 2009; M. P. Walker et al., 2002). Furthermore, adults show consolidation of order after declarative learning. Specifically, sleep promotes retention of forwards transitions between items. When participants learned sequences of three familiar words, those who slept showed greater recall for the second item in the sequence when presented with the first item. Sleeping did not benefit their recall of the first item given the second item (Drosopoulos et al., 2007). Therefore, we have some reason to believe sleep may function to strengthen forwards associations after statistical word learning, though this type of learning is generally implicit in nature.

In addition to investigating the role of sleep in memory retention and specificity, we hope to elucidate the effects of immediate testing on memory retention. Previous research has suggested that retrieving newly learned information soon after acquisition can strengthen that memory (Barcroft, 2007; Goossens, Camp, Verkoeijen, & Tabbers, 2014; Roediger & Karpicke, 2006). However, presenting potentially conflicting information during tests of recognition (in the form of low-probability strings) that take place immediately after learning could interfere with fragile, recently-formed representations. To measure the effects of immediate testing, half of our participants will complete a test immediately after the learning phase, and all return 12 hours after the learning phase to complete a delayed test. We will compare retention between the with-immediate-test and without-immediate test groups. Moreover, we will assess whether the effects of immediate testing vary depending on whether the participant sleeps or remains awake after the first session in the lab.

Lastly, a result of our design is that the learning session takes place at different times of day between our sleep and wakefulness groups. Previous studies indicate that learners may have an advantage for incidental or implicit learning tasks when the learning session occurs at a non-optimal period (Delpouve et al., 2014) or soon before a period of sleep (Himmer, Muller, Gais,

& Schonauer, 2017). Researchers have also revealed time-of-day effects in rodents, where rats (Hauber & Bareiss, 2004) and hamsters (Ruby et al., 2013) demonstrate stronger maze learning during the dark period of the day in comparison to the light period (although see Delpouve et al. [2014] for work suggesting an advantage when the learning session occurs at a non-optimal period). We plan to compare participants' performance on immediate test between the sleep condition and the wakefulness condition. If there is a learning advantage in the evening or when participants are likely to be sleeping soon, then participants in the sleep group would be more accurate during the immediate test in comparison to the wakefulness group.

4.2 Methods

We recruited a total of 107 undergraduate students at The University of Arizona to participate in this study. Before they participated, we ensured that their native language was English. All subjects received credits in their Introduction to Psychology course for participating.

4.2.1 Materials

Familiarization stimulus. For this study, we used the Set A from Experiments 1 and 2 as familiarization stimuli, though we adjusted the length of familiarization. In this study, participants listened to a language across 2 blocks of listening lasting 6 minutes each. Saffran and colleagues' (1996) study with adults implemented a longer familiarization time of 21 minutes broken into 3, 7-minute blocks. However, their language consisted of 6 trisyllabic words, while the languages in the current study contained 4 disyllabic words (reviewed below). We reasoned that a shorter familiarization time would be appropriate as our language is less complex, and a shorter exposure is less likely to cause fatigue. For this Experiment, half of the participants listened to L1 during familiarization and the other half listened to L2. Words in L1 include *pila, rodi, nibu, and dota*, while words in L2 include *lapi, diro, budo, and tani.* Neither language contains auditory cues to the location of word boundaries. Therefore, they do not contain stress contours, intonation, or pauses. The language also begins silently, ramps up in volume over the first 10 seconds, and ramps down to silence during the last 10 seconds of familiarization. This ensures that the participants cannot use the beginning or end of the familiarization stream as a segmentation cue.

Test stimulus. The test included three different types of stimulus: 4 forwards-order words (FWs), 12 part-words (PWs), and 4 backwards-order words (BW). FWs were words from the familiarization languages, and BWs were words from familiarization but with the order of the syllables reversed. For example, *pila* in L1 became *lapi* to form a BW at test. PWs consisted of the last syllable of one word and the first syllable of another; these items were made up of syllables with low sequential dependencies. The two languages were balanced so that FW items in L1 were BW items in L2 and vice versa (see Table 1) in order to ensure the form of the test items did not influence the results.

Questionnaires. All participants completed three questionnaires:: a sleep log, a general questionnaire, and the Morningness-Eveningness Questionnaire (MEQ) (Terman, Rifkin, Jacobs, & White, 2008). The questionnaires in their entirety are available in the Appendix.

Sleep Log. Participants in the wakefulness condition complete the sleep log in the morning before they listen to the artificial language. They record the time they fell asleep the night before and the time they woke up that morning. Those in the sleep condition complete the sleep log during part 2, before their delayed test, and they record their sleep and wakefulness times from the night before. Therefore, we have a measure of how long these participants slept between learning and test.

General Questionnaire. All participants complete the general questionnaire during part 2 after completing the delayed test. This form asks them to recall as many words as possible from the language they heard. Subsequent questions ask the participant to record how they felt during the study, how much they paid attention to the language, how hard they tried during the testing portion(s), and their impressions of the activities and our question. We also ask participants to record their caffeine, medication, and alcohol consumption.

MEQ. Last, participants complete the Morningness-Eveningness Questionnaire (Horne & Östberg, 1976), which provides a continuous measure of participants' preference for morning versus evening. Preference for time of day could be one reason we may see learning differences between our sleep and wakefulness groups. If our population as a whole is most alert in the evening, then perhaps these participants in the sleep group would encode more successfully than those in the wakefulness group.

4.2.2 Procedure

We assigned students to one of three delay groups: sleep (n = 33), early wakefulness (n = 35), and late wakefulness (n = 39). The sleep group listened to an artificial language between 9 and 10:30 PM. Critically, these participants slept after learning and before the retention test. The early wakefulness group listened between 9 and 10:30 AM, and the late wakefulness group listened between 10:30 AM and noon. In contrast to the sleep group, wakefulness participants remained awake for the duration of the delay between learning and test. All participants returned 12 hours after the listening session to complete a delayed test of retention. Approximately half of the participants in each group also completed a test immediately after learning. Because half of the participants in each group were tested immediately, we can a) ensure that both the sleep and wakefulness groups encoded to the same extent during familiarization and b) determine whether being tested immediately after learning influences retention of learned information.

During familiarization, participants in all groups sat quietly in a small room, listening to the artificial language through noise-canceling headphones. The language lasted approximately 12 minutes in total, and participants listened in two 6-minute blocks. Between listening blocks, subjects had a short break during which they performed an unrelated counting task.

Immediately after familiarization, half of the participants completed an immediate test. During the test, participants heard a single test item on each trial. They judged whether they recognized the item as a word from the language they heard previously; if they recognized the item, they pushed '1' on the keyboard placed in front of them. If they did not recognize the item, they pressed '9'. There were two blocks of testing, so participants heard each test item twice. The order of items was random during each block of testing. If participants had successfully learned the words, they should respond '1' to FW items and '9' to PW and BW stimuli. A researcher analyzed participants' responses to ensure no one scored zero, which would indicate that the participant reversed the key responses.

4.3 Results

For the majority of the following analyses, we use participants' average endorsement rates for each type of stimulus. Endorsement rates reflect the percentage of trials on which participants respond "yes" for each stimulus type. A successful learner will have relatively high endorsement rates for FWs and low endorsement rates for PWs and BWs.

4.3.1 Performance During Immediate Test

Figure 3 displays participants' mean endorsement rates at immediate test. To assess whether the timing of familiarization had an effect on encoding during the learning session, we began by conducting a repeated measures ANOVA with Group (sleep, early wakefulness, or late wakefulness) as a between-subjects factor and Trial Type (FWs, PWs, or BWs) as a within-subjects factor. The data for Trial Type across these conditions violates the assumption of sphericity, so we report adjusted Greenhouse-Geisser statistics. Our test of within-subjects effects revealed a Group*Trial Type interaction [$F(3.082, 77.04) = 2.696$, $p = 0.05$, $\eta_p^2 = 0.097$].

To elucidate this effect, we first analyzed the sleep group separately. We conducted a repeated-measures ANOVA with Trial Type as a within-subjects factor, finding a significant main effect of Trial Type at the immediate test [$F(2, 30) = 4.687$, $p = 0.017$, $\eta_p^2 = 0.238$]. Following up with targeted, two-tailed paired-samples t-tests, we compare the sleep group's endorsement rates of FWs to that of BWs. In the second t-test, we compared their endorsement rates of FWs to that of PWs. As we are conducting two t-tests, we use a Bonferroni-corrected alpha level of $p = 0.025$ to judge significance. The first t-test revealed that these participants, who listened to the language in the evening, exhibited significantly higher endorsement rates for FW ($M = 48.44\%$) than BWs ($M = 32.81\%$; $t(15) = -2.609$, $p = 0.020$, $d = .653$]. They also endorsed FWs at a significantly higher rate than they endorsed part-words [$M = 34.38\%$; $t(15) = -2.609$, $p = 0.020$, $d = .652$]. Thus, the participants in the sleep group endorsed the forwards words more often than part- and backwards-order words. This pattern of results suggests our participants were successful at learning and recognizing correct test items in the evening.

Next, we examined the early and wakefulness groups together in a repeated-measures ANOVA, with Group (early and late wakefulness) as our between-subjects factor and Trial Type (words, part-words, and backwards words) as the within-subjects factor. Again, our dependent variable violated the assumption of sphericity, so we report adjusted Greenhouse-Geisser statistics. We found no main effects or interactions [$F(1.382, 48.365) = 1.247$, $p > 0.05$], suggesting the early and late wakefulness groups did not differ on their performance at immediate test. Furthermore, both groups endorsed all test items at an equal rate, suggesting they did not encode syllable order (early wakefulness: M_{FW} = 43.06%, M_{PW} = 53% M_{BW} = 47.91%; late wakefulness: M_{FW} = 41.45%, M_{PW} = 44.30%, M_{BW} = 49.34%).

To sum, in our analysis of immediate test performance only, we found that the sleep group demonstrated evidence of learning, endorsing FWs at a significantly higher rate than they endorsed PWs and BWs. These participants listened to the language between 9:00 and 10:30 PM. Contrastingly, all of the individuals who listened to the language in the morning (between 9:00 AM and 12:00 PM) did not demonstrate knowledge of the word forms immediately after listening. It appears that the timing of familiarization has a significant effect on whether adults successfully learn or not as measured on a test immediately after learning.

Figure 3.

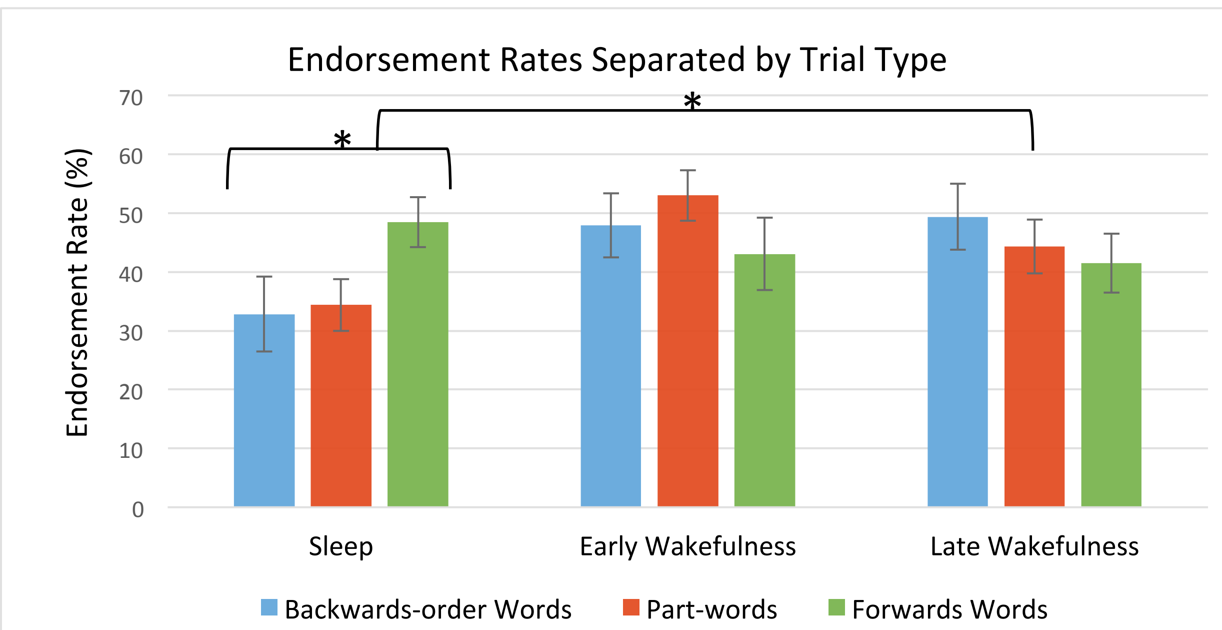

Figure 3 displays endorsement rates at the immediate test, separated by delay group and trial type. Blue bars indicate endorsement rates for backwards-order words, red bars indicate rates for part-words, and green bars indicate rates for forwards words. Error bars represent SEM. We find that only the sleep group endorses forwards words significantly more often than part- and backwards-order words, suggesting a benefit for encoding in the evening.

4.3.2 Performance During Delayed Test

To assess retention, we conducted a repeated measures ANOVA with Group (sleep, early wakefulness, or late wakefulness) and Immediate Test Group (with-immediate and without-immediate) as a between-subjects factor and Trial Type (forwards words, part-words, or backwards-order words) as a within-subjects factor, using endorsement rates at the delayed test only. There were no significant main effects or interactions [$F(3.272, 170.14) = 2.35$, $p = .065$]. Thus, all groups endorsed all test trials at an equal rate, and the immediate test did not influence performance at the delayed test.

4.3.4 Correlation Analyses

To further measure how well participants retained their knowledge of the words and the order of syllables within them, we conducted correlation analyses between their performance at the immediate test and their performance at the delayed test. We also included the MEQ in our correlation analysis in order to assess whether morningness/eveningness associated with learning or retention in any of our delay groups; a higher MEQ score reflects a preference for morning over evening. To minimize the amount of comparisons, we generated difference scores based on their endorsement rates of legal versus illegal stimuli. We used their endorsement rates to calculate difference scores because there were more part-words than forwards words presented at test. The first difference score reflected their endorsement rates for forward items minus their endorsement rates for part-words. The second difference score reflected their endorsement rates for forwards items minus their endorsement rates for backwards-order words. A positive difference score indicates participants endorsed the forwards words at a greater rate than the part- and backwards words.

First, we combined the early and late wakefulness groups as they performed equally at the immediate and delayed test. We performed correlations to assess whether immediate test performance relates to performance at the delayed test. A significant, positive correlation reflects more successful performance (high endorsement rates for words as opposed to illegal items: part-words or backwards-order words) at immediate test associated with successful performance at the delayed test. This will provide information on whether participants retained what they learned during familiarization. We calculated correlations for each delay group separately, displayed in Tables 2A and 2B.

Table 2A

SLEEP				IMMEDIATE TEST	
		Tot.Sleep (after learning)	MEQ	Part-words	Backwards words
	MEQ	.227	1	0.097	0.029
DELAYED TEST	Part-words	.213	0.002	0.288	0.39
	Backwards words	.44*	0.069	0.197	0.387

Table 2B

WAKEFULNESS			IMMEDIATE TEST	
		MEQ	Part-words	Backwards words
	MEQ	1	-.039	-0.209
DELAYED TEST	Part-words	0.143	0.569**	0.424**
	Backwards words	0.09	0.574**	0.437**

Tables 2A&B show correlations between performance at the immediate test and performance at the delayed test, separated by delay group. We included the MEQ to assess whether morningness-eveningness related to learning or retention. Table 2A shows the sleep group only, and we include Total Sleep (after learning) to assess whether the amount of sleep associated with retention in this group. Table 2B displays correlations for the combined wakefulness group (including early and late wakefulness participants). * indicates $p < 0.05$; ** indicates $p < 0.01$.

In the sleep group, we found a significant correlation of total sleep on retention of syllable order (graphed in Figure 4). The more these individuals slept after learning, the greater their difference score for FWs versus BWs. There were no significant correlations between behavior during the immediate test and the delayed test, suggesting participants who evidenced learning immediately did not necessarily continue this pattern after the delay. Considering our analysis of learning and retention in the sleep group, it appears that participants performed well at the immediate test, endorsing FWs more than BWs and PWs, but as a group they did not retain what they learned. Moreover, the MEQ was not associated with immediate nor delayed performance in this group.

Figure 4.

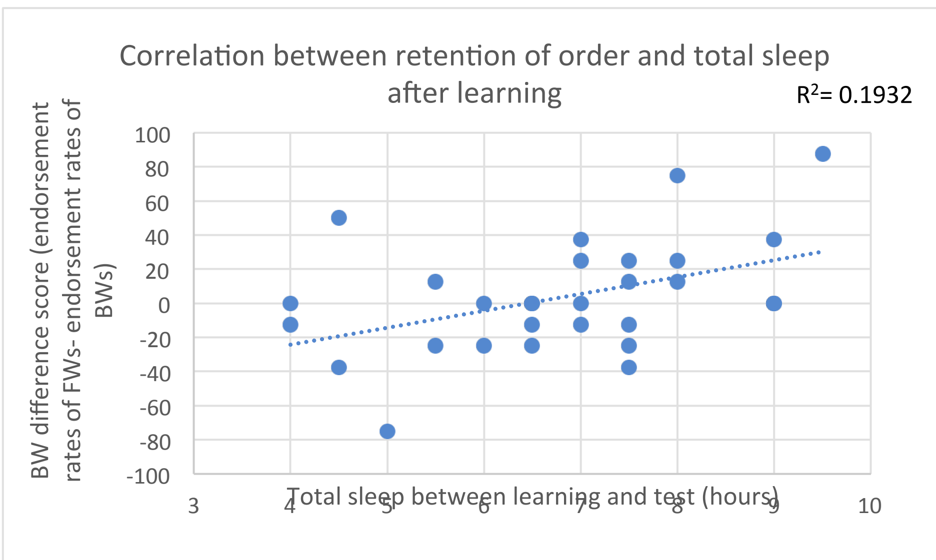

Figure 4 depicts the significant association between total sleep time after learning and retention of the order of syllables within words. While our sleep group as a whole does not demonstrate retention, we observe learning in participants who got relatively more sleep than others. This relationship indicates knowledge for order may be further consolidated with more sleep.

In the wakefulness group, immediate performance in endorsing FWs more than PWs relates to learner's ability to endorse FWs more than BWs words after the delay. At this time, we are not investigating this association further, as the relationship between these two measures is not something we are targeting theoretically. Additionally, we find evidence of retention in the wakefulness group: performance at immediate test consistently correlates with performance on the delayed test for all types of stimuli. Below, Figure 5 depicts the relationship between immediate and delayed performance on FW trials, and Figure 6 shows the relationship between immediate and delayed performance on PW trials within the late wakefulness group. These correlations indicate participants who learned successfully in the late morning retained their knowledge over the delay, continuing to endorse FWs more than PWs and BWs.

Overall, we observed a correlation between total sleep and retention of order in the sleep group. Furthermore, participants who remained awake retained what they encoded, but participants in the sleep group did not. This pattern of results indicates sleep benefits retention of order over and above encoding strength, which associates with retention across wakefulness.

Figure 5.

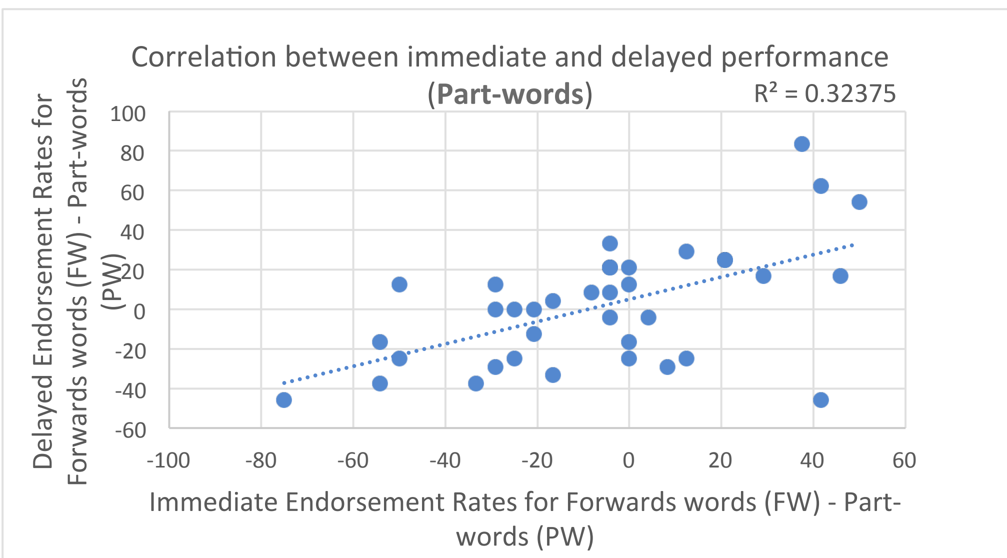

Figure 5 shows the relationship in the wakefulness group between immediate performance in endorsing FWs more than PWs and delayed performance on the same variable. This association suggests participants who endorsed FWs more often than PWs at the immediate test (indicating learning of co-occurrence frequencies) continued to do so at the delayed test, indicating retention across wakefulness.

Figure 6.

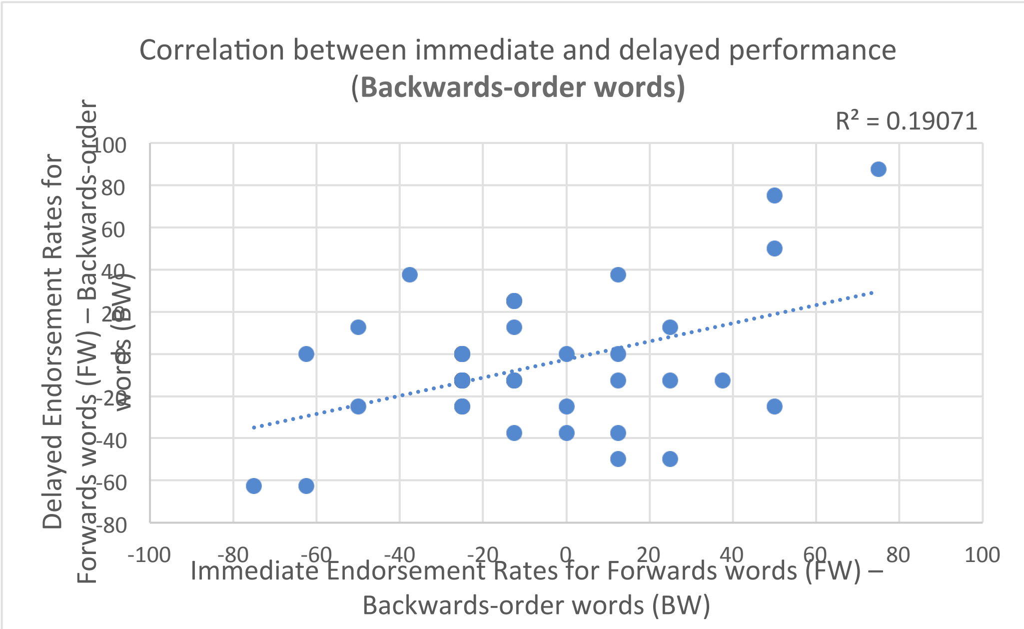

Figure 6 shows the relationship in the wakefulness group between immediate performance in endorsing FWs more than BWs and delayed performance on the same variable. This association suggests participants who endorsed FWs more often than BWs at the immediate test (indicating learning of order) continued to do so at the delayed test, indicating retention of order across wakefulness.

4.4 Discussion

We found a clear effect of time-of-day on immediate test performance, with participants in the sleep group endorsing FWs at a greater rate than they endorsed BWs and PWs. The wakefulness participants (both early and late wakefulness) did not demonstrate learning at the immediate test, endorsing all test items at an equal rate. As our sleep and wakefulness groups showed different levels of encoding, we cannot directly compare retention after sleep to retention after wakefulness. This pattern of results suggests a possible circadian influence on encoding, though the MEQ did not correlate significantly with immediate test performance in any of the groups.

Examining retention in the delay groups individually, the immediate test did not influence retention overall. Participants in all groups failed to endorse FWs more often than PWs and BWs at the delayed test. However, correlation analyses of the sleep group indicate that the participants who slept more during the delay also accepted FWs at a higher rate than BWs during the delayed test, suggesting a potential role of sleep in the consolidation of syllable order. Furthermore, in our correlation analyses of the wakefulness group, we found a positive relationship between immediate and delayed performance on PWs. Performance on BWs at the immediate test also correlated with performance on BWs at the delayed test. This indicates that while the group as a whole did not show learning, those who performed well at the immediate test continued to perform well at the delayed test, endorsing FWs more than BWs and PWs. This pattern of results suggests adults retain statistical learning across delays containing wakefulness, provided successful-enough encoding of the language before the delay.

Our finding that sleep did not confer a benefit to retention of co-occurrence frequency is in contrast to previous studies showing sleep promotes retention and generalization of grammatical rules in adulthood. For instance, sleeping after learning helped adults generalize a grammatical rule focusing on pronoun use (Batterink et al., 2014). In this study, participants needed to associate a novel pronoun with a feature of familiar objects (whether they are relatively near or far away), presented visually at the same time as the pronoun. This type of learning is quite different from the task our participants completed, in which they used sequential statistics to segment a stream into individual items. Tracking statistics to segment a string of speech may recruit different neural systems than forming associations between articles and object features presented concurrently. It could be the case that the segmentation task in our study, which does not benefit from sleep, took precedence over more general associative learning mechanisms that do consolidate learning across sleep. For instance, the hippocampus has been linked to integrating linguistic information across a conversation, including resolving an ambiguous pronoun (Kurczek et al., 2013). As hippocampal memories are consolidated with sleep, it's possible the learning in Batterink et al. (2014) was supported by the hippocampus to a greater degree than statistical segmentation, though we cannot be certain of this without neuroimaging. Nevertheless, we theorize that the differences in learning tasks between our study and Batterink et al. (2014) drive the contrasting results regarding sleep-dependent memory consolidation. While some types of statistical learning benefit from sleep, such as the task used by Batterink et al. (2014), the current study suggests representations gained through statistical *word segmentation* are not consolidated with sleep.

Additionally, Batterink and Paller (2017) required participants to form phrases by putting novel words in order. Unbeknownst to participants, the words were separated into 6 different classes, and each class could only occur in specific positions in each phrase. Furthermore, some classes had high co-occurrence frequency with other classes (e.g. class A and class B occurred in sequence often) while others had relatively low co-occurrence frequency (e.g. class A and class C occurred in sequence occasionally but not often). After learning through feedback, participants slept, then later completed a test in which they had to correctly order the words into phrases that had not appeared during learning. The researchers presented an audio clip of the grammar for half of the participants while they slept to promote memory consolidation. After sleeping, all participants successfully ordered the words to generate a majority of the test phrases, though those who incurred targeted memory reaction during sleep showed enhanced performance at test. Critically, all participants' test performance was not influenced by the co-occurrence frequency between classes. They were equally accurate at ordering test sentences that contained classes with high co-occurrence frequencies and those with low co-occurrence frequencies, and targeted memory reactivation did not benefit retention of high-frequency sequences of classes versus low-frequency sequences of classes. This finding suggests active processes during sleep do not benefit rote memory for co-occurrence frequency, which is similar to our results with adults, in which sleep does not benefit participants' retention of co-occurrence between syllables. However, both studies suggest participants retain the general order of linguistic constituents with sleep (e.g. item A comes before item B), allowing participants to properly order word classes to generate new phrases (Batterink & Paller, 2017). Additionally, in Experiment 3 we found a correlation between total sleep time and participants' ability to correctly reject backwards-order words in our study (displayed in Figure 3), linking sleep to successful retention of order but not co-occurrence frequency.

Overall, we offer correlational evidence that sleeping after statistical word learning promotes memory for the order of syllables within words. However, participants who stayed awake after learning in the morning retained their knowledge across the delay, suggesting sleep may not be necessary for retention of statistically learned words and the order of syllables within them.

Chapter 5: Experiment 4

Retention of Nonadjacent Dependencies: Does Sleep Promote Veridical Memory in 18-month-old Infants?

5.1 Introduction

As discussed in Experiments 1 and 2, the ability to segment words from speech relies on tracking conditional probabilities between adjacent syllables. However, language does not only contain relationships between adjacent items. It is important to note that they also typically include long-distance dependencies, in which two constituents are related but separated by one or more words. For example, in English there a verb phrases such as *he is running, the girl is laughing,* and *the dog is barking.* In all of these phrases, the verb *is* and the morpheme *-ing* remain constant while the rest of the vocabulary varies. In fact, any regular verb can occur between *is* and *-ing* and the phrase will remain grammatical. These two constituents form a nonadjacent dependency; they are consistently paired yet they are separated by at least one intervening element.

Studies assessing acquisition of nonadjacent dependencies in grammar often use artificial languages that contain phrases with the forms aXb and cXd. In these languages, *a, b, c,* and *d* are single words, such as *pel, vot, rud,* and *jic.* The X items vary and are drawn from a pool of disyllabic words. Research indicates that listeners can learn the two nonadjacent dependencies in this grammar when there are at least 18-24 unique X items included in the corpus. It appears that high variability of the medial item promotes learners to realize the consistency of the nonadjacent elements and form an association between them (Gómez, 2002; Gomez & Maye, 2005). While this task is more difficult than word segmentation, as the related constituents in a nonadjacent dependency occur remotely from one another, Gómez and Maye (2005) revealed that infants as young as 15 months of age are able to learn the nonadjacent dependencies while listening to an artificial grammar. They can discriminate the grammar they listen to during familiarization (e.g. aXb, cXd) from a grammar that violates the nonadjacent dependency (e.g. aXd, cXb).

Following this line of research and considering the role of sleep for memory retention in adults, Gómez, Bootzin, and Nadel (2006) asked whether 15-month-old infants could remember two nonadjacent dependencies across a delay of four hours. They also manipulated whether infants slept soon after listening to an artificial grammar in order to assess whether sleep benefitted their memory. In this study, all participants listened to a grammar containing two NADs (e.g. *pel-X-jic* and *vot-X-rud*). Half of the infants napped between learning and test while the other half remained awake. Four hours after listening to the grammar, all infants completed a test using the Head-Turn Preference procedure (Kemler Nelson et al., 1995).

The authors analyzed the infants' listening time in two different ways. First, they assessed whether the infants showed specific, veridical memory for the NADs and the vocabulary that comprises them. To obtain this measure, the authors calculated the average listening time to familiar trials – those consistent with the grammar they heard before the delay – along with the average listening time to unfamiliar trials, which violated the NADs in the grammar they listened to previously. They subtracted the average listening to unfamiliar trials from the average listening time to familiar trials. With this measure, a difference score varying significantly from 0 was an indication that infants retained the exact NADs they had heard during familiarization before the delay. Secondly, the authors were interested in whether infants would be influenced by the identity of the first trial. They reasoned that if participants retained the nonadjacent dependencies at an abstract level but forgot the specific vocabulary that comprised them, then the infants would hear the vocabulary used in the first test trial and use this information to guide their behavior for the remainder of the test. More specifically, the infants would have remembered that there is a relationship between the first and third item in each phrase but not the vocabulary that fills those slots. During the first test trial, they hear the vocabulary used and place those items into their nonadjacent dependency slots. Then, they base their behavior during the test on the nonadjacent dependencies in the first trial, discriminating trials that are consistent with this grammar from trials that are inconsistent with this grammar.

Overall, Gómez et al. (2006) found that 15-month-old infants who stayed awake could discriminate the familiar grammar, heard before the delay, from an unfamiliar grammar that violated the NADs they learned previously. The infants who napped during the delay did not discriminate the familiar grammar from the unfamiliar grammar. These infants listened equally to both familiar and unfamiliar trials, indicating that they did not remember the specific NADs they heard before the delay. While these infants did not show veridical memory for the NADs, they did show the abstraction effect. These infants listened longer to trials that were consistent with the identity of the first trial. This indicates that these infants remembered that there is a relationship between the first and third items in each phrase but did not remember which vocabulary words formed the nonadjacent dependencies. Thus, it appears that 15-month-olds can retain the NADs over wakefulness, but sleep promoted forgetting of the specific vocabulary that made up the NADs.

In a follow-up study, (Hupbach et al., 2009) questioned whether napping soon after learning would benefit 15-month-olds' retention of NADs over a delay of 24 hours. They replicated the methods from Gómez et al. (2006), but all infants were tested 24 hours after listening to the artificial language. Half of the infants stayed awake in the four hours after listening and half took a nap in this time. Now, infants who did not sleep after exposure to the artificial language were not able to discriminate the familiar grammar from the unfamiliar, and they also did not show a preference for trials that were consistent with the grammar in the first test trial. This result indicates that infants who stay awake after learning do not remember the NADs in the familiarization language, and they do not retain abstract knowledge of NADs that would allow them to recognize the pattern in the first test trial. Infants who napped, however, continue to show the first-trial effect: they listen longer to trials that are consistent with the grammar in the first test trial. These infants retained their abstract knowledge of NADs across a delay of 24 hours. Taking these results together with those of Gómez et al. (2006), the authors concluded that a nap after learning is necessary for 15-month-old infants to retain anything from their brief exposure to the NADs during familiarization.

Considering neural systems take time to develop, and the role of sleep in memory retention may change with age, perhaps older children will show increased specificity in their representations of NADs. To investigate this question further, we use the same stimuli and procedure as Gómez et al. (2006) but recruit 18-month-old infants. As these infants are further along in development and have more mature neural systems than 15-month-olds, it is possible we may see different effects of sleep on retention in these older infants. Additionally, Santelmann & Jusczyk (1998) demonstrated that 18-month-old infants can discriminate legal NADs in English (e.g. *is running*) from illegal NADs in English (e.g. *can running*), suggesting these infants are more mature in their knowledge of NADs in general. Therefore, we may expect more robust encoding during familiarization in these infants, which could result in greater benefit from sleep.

5.2 Methods

We recruited 29 families with 18-month-old infants (*M* = 18.18 months, range = 17.53 – 18.82 months) to participate in this study. Before scheduling, we ensured that the infants had a relatively regular nap time. We considered an infant's nap to be regular if it typically fell into a two-hour window (e.g. if an infant napped between 10 AM and noon). Of the families who participated, we excluded the data from infants in our analyses if a parent or sibling interfered during familiarization (6) or test, if the infant did not participate in the test session or became fussy and quit attending during the test (7) in the case of deviations from protocol/experimenter error (10) or due to technical errors (4).

We assigned participants either to a nap (N = 21) or wakefulness condition (N = 8). We scheduled infants in the nap condition so that familiarization took place approximately 1 hour before their expected nap time. After familiarization, these infants took a nap as usual, and four hours after familiarization the families came to the lab to be tested using the Head-Turn-Preference Procedure (Kemler Nelson et al., 1995). Infants in the wakefulness condition listened to the familiarization language at a time when they were expected to stay awake during the following four hours. After the four-hour delay containing only wakefulness, the infants came to the lab to complete the test.

5.2.1 Materials

Familiarization stimuli. We adapted our stimuli from Gómez et al., (2006). During familiarization infants listened to one of two grammars, each containing two nonadjacent dependencies. Phrases in the first grammar (G1) consisted of phrases with the structure *pel-X-jic* and *vot-X-rud*, in which the first element and third element were consistently paired. The second grammar (G2) contained phrases following a contrasting pattern, *pel-X-rud* and *vot-X-jic.* Thus, the phrases from G1 violated the grammar in G2 and vice versa. In both grammars, the medial X elements were drawn from a bank of 24 disyllabic words. These included: *wadim, kicey, puser, fengle, coomo, loga, gople, taspu, hiftam, deecha, vamey, skiger, benez, gensim, feenam, laeljeen, chila, roosa, plizet, balip, malsig, suleb, nilbo, and wiffle.* Both grammars contained 48 unique phrases that were each repeated 5 times, and phrases were separated by 1 second of silence during familiarization. The number of exposures was selected to ensure ample time for encoding (following Gomez et al., 2006), though researchers have not empirically tested how exposure time influences learning of NADs. Familiarization lasted a total of 14 minutes, ending after 5 repetitions of each phrase.

Test stimuli. Half (8) of the test trials contained phrases from one grammar (G1 or G2). The remaining half of the trials contained phrases from the other grammar. Each trial contained 6 phrases and lasted for a maximum of 17 seconds. Critically, half of the infants heard phrases from G1 on the first trial and the other half listened to G2 on the first trial. Following the findings of Gomez et al. (2006) we implemented this manipulation so that we can assess whether the grammar in the first trial influences infants' behavior for the remainder of the test.

5.2.2 Procedure

Familiarization. All infants listened to the familiarization language in a quiet room in their home accompanied by their caregiver and the researcher. Both the caregiver and researcher refrained from speaking while listening, and they entertained the infant silently to keep them happy and engaged.

Test. We tested infants using the Head-Turn-Preference Procedure (Kemler Nelson et al., 1995). As in Gomez et al., (2006), the researcher and caregivers wore noise canceling headphones throughout the test, and the researcher asked caregivers to avoid interacting or communicating with their child to prevent interference.

5.3 Results

5.3.1 Veridical Memory

To assess their veridical memory for the nonadjacent dependencies, we first calculated infants' average listening time to familiar and unfamiliar trials. For instance, if infants listened to G1 before the delay, then test trials containing phrases from G1 would be familiar and trials with G2 would be unfamiliar to the infant. For the nap and wakefulness groups separately, we subtracted their looking time to unfamiliar trials from their looking time to familiar trials. If infants remember the grammar they listened to before the delay, they would distinguish the familiar from the unfamiliar grammar, and their difference score would vary significantly from zero.

Before beginning our primary analyses, we first ensured that neither the familiarization grammar (G1 or G2) nor gender influenced infants' difference scores at test. For the nap and wakefulness groups separately, we conducted a t-test comparing performance of infants who heard G1 to infants who heard G2, finding no significant difference in the nap group [M_{G1} = 1.28s, M_{G2} = 1.32s; $t(19) = -0.04$, $p > 0.5$], nor the wakefulness group [M_{G1} = 0.72s, M_{G2} = 1.59s; $t(6) = -0.691$, $p > 0.05$]. Next, we conducted a t-test comparing the males' difference scores to the females' scores. There was no effect of gender in the nap group [M_{males} = 1.10s, M_{females} = 1.58s; $t(19) = -0.507$, $p > 0.05$]. There was a significant effect of gender in the wakefulness group though our sample size is currently small [$t(6) = 2.82$s, $p = 0.03$], with males (M = 2.56s) showing larger difference scores than females (M = 0.14s). It is important to note that we currently have 8 infants in the wakefulness group, with 3 females and 5 males. When we resume data collection after the coronavirus pandemic, we will reanalyze to assess whether the effect of gender persists once we have more statistical power. For the following analyses, we collapsed across these variables for the remaining analyses of veridical memory. Current looking time differences for both groups are displayed in Figure 7.

We conducted one-sample t-tests, comparing the nap and wakefulness groups' difference scores to 0. Difference scores were significantly greater than 0 in the nap condition ($M = 1.31$s; $t(20) = 2.845$ $p = 0.010$, $d = 0.621$], with longer looking during familiar trials ($M = 8.11$s) than unfamiliar trials ($M = 6.80$s). Infants in the wakefulness condition did not discriminate the familiar grammar from the unfamiliar, as their difference scores did not differ significantly from 0 [$M = 1.05$s; $t(7) = 1.797$, $p > 0.05$]. We continued by comparing veridical retention between the nap and wakefulness groups using a two-tailed, independent samples t-test. We found no significant difference between the groups [$t(27) = -0.306$, $p > 0.05$]. Though the wakefulness infants' difference scores do not significantly exceed 0, their scores are not significantly lower than those of the nap infants. With sufficient sample size we expect infants in the wakefulness group to discriminate the familiar from unfamiliar grammars. We predict they will show veridical retention given that 15-month-olds retained specific representations across a similar period of wakefulness (Gomez et al., 2006).

Figure 7.

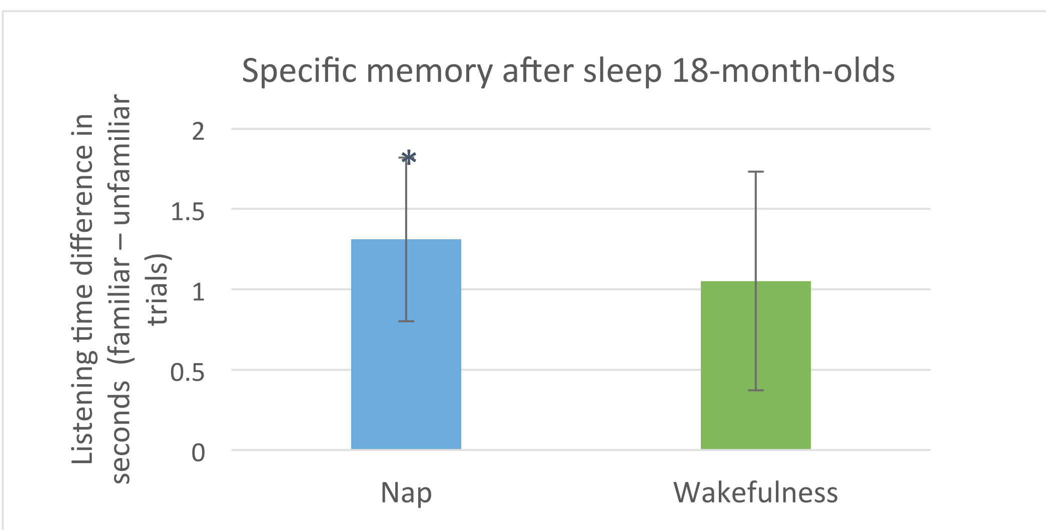

Figure 7 displays the listening time difference in seconds to trials familiar to the infant minus listening time to unfamiliar trials. Error bars reflect SEM. The nap group, in blue, shows a difference score significantly greater than 0, indicating they listened significantly longer to familiar trials. At this point in data collection, the wakefulness

5.3.2 Generalization

Following Gomez et al. (2006), we also assessed whether infants retained an abstract representation of the nonadjacent dependency that facilitates generalization to grammars similar to that heard during familiarization. To measure generalization, we noted which grammar each infant heard during the first test trial. We calculated (1) the average looking time to trials consistent with the first trial and (2) the average looking time to trials inconsistent with the first trial. For example, if infants heard G1 on the first trial, we included all G1 trials (including the first trial) in our average looking to trials consistent with the first trial. For the same infants, data from all G2 trials was included in our average looking to trials inconsistent with the first trial. We calculated a difference score by taking the average looking time to consistent trials and subtracting the average looking time to inconsistent trials. If infants have nonspecific, abstract representations of the nonadjacent dependencies from the familiarization grammar, we would observe a difference score that varies significantly from 0. This pattern of results would indicate that infants remember there is a relationship between the first and third items in each phrase, but they do not remember the specific vocabulary that forms those relationships. During the first test trial, they hear the vocabulary in those nonadjacent dependencies, and use this information to guide their looking for the remainder of the test.

Before beginning our primary analyses, we first assessed whether there was an effect of gender or familiarization grammar on generalization. First, in the nap group we conducted t-tests comparing males' difference scores ($M = 0.75$s) to females'($M = 0.83$s), finding no difference based on gender [$t(19) = -0.077$, $p > 0.05$]. We also found no main effect of grammar [$t(19) = -1.313$, $p > 0.05$]. In the wakefulness group, we conducted t-tests comparing performance between genders and familiarization grammars. We found an effect of gender

[$t(6) = 4.942$, $p = .003$, $d = 3.609$], with males showing greater difference scores ($M = 2.56$s) than females ($M = -0.794$s). However, due to our current sample size in the wakefulness group (n = 8) this effect may not be representative. We found no effect of grammar.

To conduct our main analyses, we used two, two-tailed t-tests comparing the difference scores to 0: one for the wakefulness group and one for the nap group. We found that neither the nap group's average [M = 0.788s, $t(20) = 1.554$, $p > 0.05$] nor the wakefulness group's average [$M = 0.46$s, $t(7) = 0.677$, $p > 0.05$] differed significantly from 0. We then compared the nap group to the wakefulness group using a two-tailed, independent samples t-test, finding no difference in generalization difference scores between groups [$t(27) = -0.35$, $p > 0.05$]. Neither the nap nor the wakefulness infants generalized their knowledge of NADs to the first test trial.

5.4 Discussion

Overall, we found that the infants who napped retained veridical representations of the NADs they learned during familiarization. Our small sample of wakefulness infants did not discriminate the familiar from unfamiliar NADs, though we expect them to do so once we have sufficient power to detect an effect. We hold this hypothesis because the 15-month-old infants in Gomez et al. (2006) successfully retained the NADs over wakefulness. Furthermore, our current group of 8 infants is behaving in line with this prediction, looking approximately 1 second longer to familiar trials than unfamiliar trials. However, their current difference scores do not vary significantly from 0, and we will resume data collection once possible. At this time, we cannot make definite claims on the effects of sleep versus wakefulness on veridical retention at 18 months of age.

Looking at our measure of generalization, we found an effect of gender in our wakefulness group, though we cannot be certain this is a representative effect due to our small sample size. Overall, neither group as a whole was biased by the identity of the first trial at test. In combination with our results of veridical memory, this pattern of results suggests infants retain veridical representations of the NADs from familiarization across a delay of four hours containing a nap. Because infants remembered the exact NADs from familiarization, they were able to distinguish these from NADs that violated the pattern heard before their nap; the identity of the first test trial did not matter.

Comparing our results to those of Gomez et al. (2006), it appears that there is a development in the effect of sleep on memory between 15 and 18 months of age. Using the exact same stimuli and procedures, Gomez and colleagues found that 15-month-olds fail to demonstrate veridical retention of the NADs they learned before they slept. Rather, they retained a general knowledge of NAD structure, and they recognized the NADs in the first test trial regardless of whether they were consistent with the pattern in their familiarization language. The authors reason that these infants forgot the specific vocabulary making up the NADs but remembered the predictive relationship between the first and third words in each phrase. At test, the 15-month-olds hear the NADs in the first test trial and place the corresponding vocabulary in their NAD "slots." They go on to use these NADs to guide their behavior for the remainder of the test. At 18 months, after sleeping infants are able to discriminate the exact NADs they heard during familiarization from a grammar that violates those NADs. Thus, by 18 months of age, it appears that sleep may strengthen the veridical representations of the NADs from the familiarization grammar.

Older children might see more of a benefit from sleep because they are more advanced learners. Their general knowledge of NADs in English might allow them to encode more effectively or form more robust veridical representations during exposure to an artificial language. Alternatively, this window in development could be a time in which neural systems supporting sleep-dependent memory consolidation reach sufficient maturity to support stabilization of detail.

We observed that 18-month-olds are not biased by the identity of the first test trial when the testing grammars use the same vocabulary as the familiarization grammar. Their lack of generalization likely stems from their strong veridical memory for the familiarization grammar. Generalization is important in acquisition of grammar, as a language user needs to be able to recognize and understand grammatical sentences that they have not heard before. In Experiment 5, we continued to assess whether 18-month-olds have a flexible enough memory to generalize their knowledge of NADs to a grammar with completely novel vocabulary. Furthermore, we asked whether sleep offers a benefit to retention over wakefulness.

Chapter 6: Experiment 5

Does Sleep Promote Generalization of Grammatical Patterns in 18-Month Old Infants?

6.1 Introduction

In Experiment 4, 18-month-olds who nap after learning retained a veridical representation of the artificial language. While we are still collecting data from infants who remain awake after learning, we expect these participants to behave similarly by demonstrating specific knowledge of the NADs and the vocabulary that comprise them. To extend the findings of Experiment 4, we asked if 18-month-old infants will generalize their knowledge of NADs if the testing conditions are novel enough. In the previous experiment we used the same vocabulary between familiarization and test, and 18-month-olds infants have a detailed enough memory to discriminate the familiar grammar from the unfamiliar grammar. In Experiment 5, we ask whether 18-month-old infants will be able to recognize the NAD pattern at test when the familiarization grammar uses different vocabulary from the testing grammars. As in Experiment 4, one group of infants stayed awake in the four hours after learning while the other group napped after learning. This design will allow us to assess whether sleep promotes abstraction of the pattern over an equal period of wakefulness.

There are several studies demonstrating the benefits of sleep for linguistic generalization in addition to Gomez et al. (2006) and Hupbach et al. (2009). As reviewed, Batterink and Paller (2017) used targeted memory reactivation during to increase generalization of a grammatical rule, demonstrating sleep's active role in rule abstraction. Even without targeted memory reactivation, sleep promotes grammatical generalization (Batterink et al., 2014). Furthermore, Friedrich and colleagues (2015) showed that napping benefits generalization in 9-month-olds. In this study, infants who napped after learning labels for novel categories extended the labels to novel exemplars, but infants who stayed awake after learning did not (Friedrich et al., 2015).

How might sleep be promoting generalization? Lewis and Durrant (2011) propose that repeated reactivation of memories during sleep play an important role in abstraction. Namely, when there are commonalities between memories, repeated reactivation encourages the consistent elements to be particularly strengthened as they are more greatly activated than elements that correspond to only one memory. Particularly in development, sleep may also function to actively weaken unnecessary connections formed during learning while also strengthening maintained connections (Li et al., 2017). Along these lines, Tononi and Cirelli (2006) propose that processes during sleep function to decrease the synaptic load by pruning weaker connections, called synaptic downscaling. If the specifics of a memory are not highly encoded, then the irrelevant details will be pruned away with sleep, leaving only the most stable aspects of the memory.

In the case of NADs, sleep could promote forgetting of the vocabulary, but infants retain the relationship between the first and third words as this was the most consistent aspect of the familiarization grammar. More specifically, 15-month-olds may separately encode positional information at anchor-points (i.e. the vocabulary at the beginning and end of each phrase) while also encoding the NAD relationship between these items (that the constituents in the 1st and 3rd positions are related). In other words, infants form two levels of representation during learning. They encode the exact words that occur at the beginning and ends of phrases (Benavides□Varela & Mehler, 2015) and learn the association between these specific items, but they also separately represent the existence of a relationship between the item in the first position and the item in the third position. Their representation of the exact words is not strong enough to survive synaptic downscaling because the vocabulary varied to some extent during familiarization (e.g. the first position could contain either “pel” or “vot’). Contrastingly, the general NAD relationship between the 1st and 3rd items was consistent across all phrases in the grammar, so infants may have encoded this pattern to a greater extent than they encoded the exact vocabulary. Thus, the general NAD relationship survives downscaling but the specific vocabulary making up the NADs does not, resulting in generalization after sleep.

Eighteen-month-olds, on the other hand, form stronger representations during encoding because of their further development (Gomez, 2002; Gomez & Maye, 2005; Santelmann & Jucszyk, 1998). Like 15-month-olds, they may separately encode positional information (which vocabulary occurs at the beginning and end of each phrase and how words associate) along with the NAD relationship, that the 1st and 3rd items are associated. Stronger encoding in this age group generates connections that persist after synaptic downscaling, so 18-month-olds show veridical retention across sleep rather than generalization (Experiment 4). We hypothesize that their representation of the general NAD rule would also persist against downscaling across sleep, as in 15-month-olds (Gomez et al., 2006). If this is the case, 18-month-olds should apply this rule to similarly-structured phrases that contain novel vocabulary, as their NAD representation does not include specific lexical representations: these are stored separately from the general rule.

In this Experiment, we measured 18-month old’s ability to generalize their knowledge of NADs to completely novel vocabulary, along with whether sleep provides a benefit to retention of the NAD rule. Infants first listened to an artificial grammar containing two NADs. Then, half of the participants slept while the other half remained awake in the four hours after listening. To test the infants, we presented artificial grammars with completely novel vocabulary making up the NADs. If infants have a general, abstract representation of the NADs heard before the delay, then they would tune in to the NADs present in the first test trial and use them to guide their behavior for the remainder of the test session.

6.2 Methods

We recruited 33 families with 18-month-old infants (*M* = 18.33 months, range = 17.62 – 19.92 months) to participate in this study. All families were monolingual English speaking, and none of the infants' immediate family members received speech therapy.

We randomly assigned 17 infants to the nap condition and 16 to the wakefulness condition. As in Experiment 4, we scheduled familiarization so that infants in the nap condition were expected to nap soon after listening to the artificial language. Researchers scheduled the wakefulness infants so that familiarization took place before a period of time in which the infants were expected to stay awake for at least 4 hours. All infants came to the lab four hours after familiarization to complete the test using The Head Turn Preference Procedure (Kemler Nelson, 1995).

6.2.1 Materials

Familiarization Stimuli. We created a single familiarization grammar that followed the structure of those in Gomez et al. (2006). It contained two NADs: *guf*-Y-*zam* and a *miv*-Y-*fop*. The Y items came from a novel set of 24 unique words: *burkid, drakel, fraggle, fupred, gaffin, hager, harnet, hilgub, kondit, labo, lotad, mabbit, minka, nutter, plicka, pumraz, quabit, rezzit, shodo, stenga, triftan, vozlet, wahav, and zeenel*. The grammar was made up of 48 different phrases. Each phrase repeated 5 times separated by 1 second of silence. The grammar played for approximately 14 minutes, and all infants listened to the same grammar.

Test Stimuli. We used G1 and G2 test stimuli from Gomez et al. (2006) and Experiment 4 as the test stimuli. We will refer to them as Test Grammar 1 (TG1) and Test Grammar 2 (TG2) for this Experiment. To remind you, TG1 contained phrases that followed the pattern *pel-X-jic* and *vot-X-rud.* TG2 followed a contrasting NAD pattern: *vot-X-jic and pel-X-rud.* The X items were drawn from a bank of 24 disyllabic words, different from the familiarization grammar used in this experiment. Half (8) of the test trials contained phrases from TG1, the other half contained phrases from TG2. Furthermore, we counterbalanced whether infants heard TG1 or TG2 on the first test trial.

6.2.2 Procedure

Familiarization. As in Experiment 4, familiarization took place at the participants' homes. Infants listened to the familiarization language in a quiet room in their home accompanied by their caregiver and the researcher. An assistant accompanied the researcher to play with any siblings in a separate room during familiarization. Both the caregiver and researcher refrained from speaking while listening, and they entertained the infant silently to keep them happy and engaged.

Test. The testing procedure was identical to that in Experiment 4 (with the exception of the different test stimuli). Infants completed the test in Head-Turn-Preference Procedure (Kemler Nelson et al., 1995). As in Gomez et al. (2006), the researcher and caregivers wore noise canceling headphones throughout the test, and the researcher asked caregivers to avoid interacting or communicating with their child to prevent interference with testing procedure.

6.3 Results

As we only presented novel vocabulary at test, we generated one difference score per infant, taking their average looking to trials consistent with the first test trial and subtracting average looking to trials inconsistent with the first trial type. For example, if infants heard TG1 first, we would calculate [average TG1 looking] – [average TG2 looking], and vice versa if they heard TG2 first. In order to ensure infants' behavior was not influenced by their gender, we conducted independent-samples t-tests comparing the males' difference scores to the females'. There was no difference in performance based on gender for the nap group [$t(15) = 0.584$, $p > 0.05$], nor for the wakefulness group [$t(14) = -0.745$, $p > 0.05$] so we collapsed across this variable for the remaining analysis. Figure 8 displays the difference scores of the nap and wakefulness groups.

Then, we conducted two, two-tailed, one-sample t-tests. The first test compared the nap group's difference score to 0, and the second compared the wakefulness group's difference score to 0. A score significantly different from 0 indicates that infants noted the NADs in the first test trial and discriminated those NADs from the contrasting grammar. In the nap group, there was a significant difference between the infants' difference scores and 0 [$M = 0.98$s; $t(16) = 2.926$, $p = 0.01$, $d = 0.71$] indicating they generalized their knowledge of NADs to the grammar presented in the first test trial. The wakefulness group's difference scores did not significantly differ from 0 [$M = 0.15$; $t(15) = 0.213$, $p > 0.05$]. They did not extend their knowledge of NADs to recognize the ones present in the first test trial. We then compared the nap and wakefulness groups using a two-tailed, independent samples t-test comparing their difference scores. We found no difference between the groups [$t(31) = -1.096$, $p > 0.05$]. Though the nap group generalized while the wakefulness group did not, we likely find no difference between them due to the relatively large variability with the wakefulness group in particular.

Figure 8.

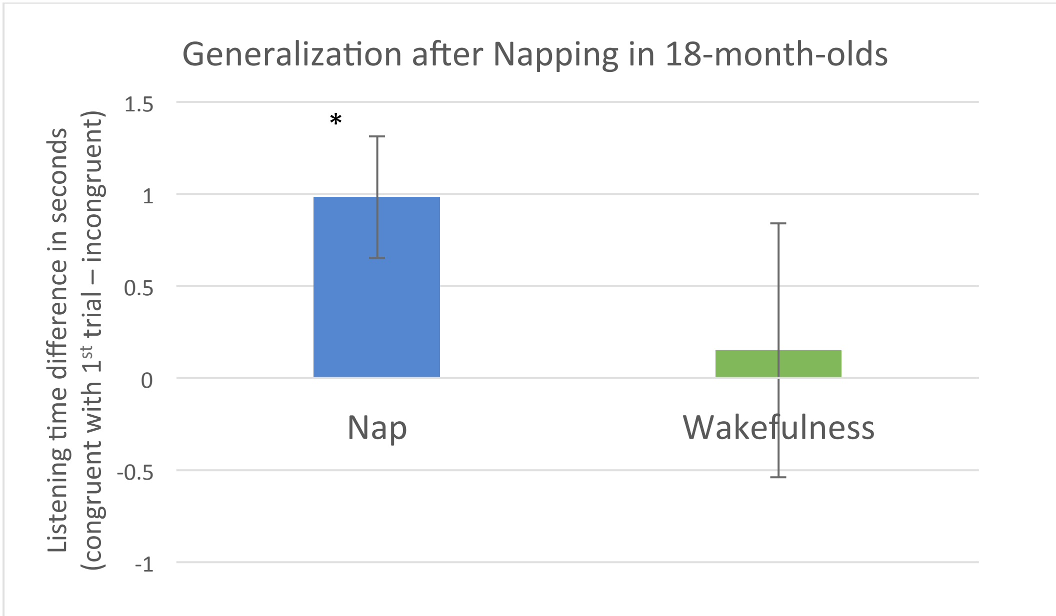

Figure 8 depicts our measure of generalization: comparing listening time to trials congruent with the first trial type versus listening time to incongruent trials. Error bars reflect SEM. Only the nap group showed a generalization effect; the wakefulness groups' difference score did not vary significantly from 0.

6.4 Discussion

What processes are likely to support 18-month-olds' generalization to novel vocabulary across sleep? We propose that sleep promotes retention at two levels of representation in older infants. Namely, sleep benefits detailed, veridical representations of the NADs (Experiment 4). They remember which words occurred in the 1st and 3rd positions in each phrase and the relationships between these specific items (e.g. that "vot" was paired with "jic" and not "rud.") unlike at 15 months. In tandem, sleep promotes retention of the general NAD rule itself: knowledge of a relationship between first and third words in each phrase. With sleep, 18-month-old infants retain the exact NADs they learned earlier and *also* more general, flexible knowledge of the NAD relationship, allowing them to recognize NADs in phrases with completely novel vocabulary.

We also found that 18-month-olds do not generalize to new grammar over wakefulness, though we propose they encode the general NAD relationship between the 1^{st} and 3^{rd} items during learning. This finding indicates the importance of sleeping after learning for retention of the abstract NAD relationship in particular, as infants do not retain this information across wakefulness. In Experiment 4, 18-month-olds retained the exact NADs heard from familiarization. Why would they retain the veridical representations of the NADs but not retain the abstract relationship between the 1^{st} and 3^{rd} slot of each phrase? We suggest that this flexible representation is more fragile and more difficult to stabilize over wakefulness, and sleep is necessary for stabilization. It may be the case that the abstract representation, which allows infants to recognize the NAD relationship in new vocabulary, is more prone to interference than their veridical memory for the NADs from familiarization. The abstract representation may decay more quickly across waking experience than the specific memory for the familiarization NADs. Therefore, at both 15 and 18 months, infants who do not nap after learning will show veridical retention only (Gomez et al., 2006; Experiment 4 in this dissertation). With sleep, infants at both ages retain a general predictive relationship between the first and third items in each phrase that does not include the specific vocabulary filling the NAD slots, facilitating generalization (Gomez et al., 2006; Experiment 5 in this dissertation). At 18 months, however, sleep also promotes stabilization of detail, so older infants who nap show veridical retention (Experiment 4) in addition to generalization (Experiment 5). In the general discussion, we consider what factors are driving the differences between sleep-dependent memory consolidation from 15 to 18 months of age.

Related to sleep's roll in stabilization and generalization, there is some evidence that sleep promotes stabilization in infancy so that memories persist over a longer delay. Namely, Hupbach and colleagues (2009) observed that without a nap after learning, 15-month-olds do not retain anything about the familiarization grammar 24 hours later. Contrastingly, infants who napped soon after learning continued to remember the abstract relationship between the 1st and 3rd word in each phrase after a 24-hour delay. This pattern suggests sleep strengthened or stabilized the abstract memory for the NAD relationship, so it was longer lasting than what infants retained when they did not sleep. Furthermore, in regard to generalization there may be active processes occurring during sleep that promote long-term retention of the general dependency between the first and final items in each phrase of the artificial grammar. For example, Lewis & Durrant (2011) propose that reactivations of recently encoded memories encourage integration across exemplars. Therefore, as the veridical representations of the familiarization NADs are activated, the system further extracts or strengthens the common general dependency between the first and third items.

Alternatively (or in concert), synaptic homeostasis might selectively retain the connections representing the general NAD rule by eliminating connections related to irrelevant information. Such unimportant details could include aspects of the input not as necessary for the grammatical structure, such as the speaker's voice, the intonation patterns of each phrase, and potentially the specific representation of the 24 individual medial items. These aspects of the grammar are not vital to retention of the exact NADs heard during familiarization, and they do not benefit an abstract representation of the NAD rule (that the 1st and 3rd word in each phrase are associated). Without sleep, maintaining the strength of the abstract representation may not be possible, as new connections are formed during wakefulness. New learning during waking experience could cause the existing representation of the NAD rule to weaken, especially while infants extract other general patterns from their environment (McCloskey & Cohen, 1989). This abstract information decays to a greater extent than their memory for the specific NADs because the abstract memory requires the infant to code and retain the relationship completely separately from individual words, making it less concrete and potentially less stable.

6.4.1 Benefits of Generalization

How beneficial is an abstract, flexible memory for grammar? English contains NADs like "she is laughing," where "is" and "-ing" are consistently paired. English also has phrases like "he was laughing," which violates the dependency between "is" and "-ing." If infants maintained a rigid memory for "is-X-ing," they would not accept "was-X-ing" as a grammatical phrase. Furthermore, advanced language users can generate completely novel sentences; flexible representations of grammatical rules allow listeners to recognize grammatical sentences even when some vocabulary is novel. Thus, some level of flexibility is important. However, infants must form specific enough memories to realize when a nonadjacent dependency is violated, like in phrases such as "can-X-ing." Therefore, learners must form a balance between veridical and flexible representations of grammatical rules.

In this study, we presented nonadjacent dependencies using completely novel vocabulary from familiarization, and infants mapped their knowledge onto these novel phrases. This may not be an exact mirror of natural language processing, since morphemes making up nonadjacent dependencies are closed classes with small sets. In other words, in natural language acquisition, once an infant learns the frequent nonadjacent dependencies relevant for their language, they do not need to continue extending the NAD relationship to similar strings. However, 18-months is a pivotal time in language acquisition, it could be the case that infants are still learning about the

NADs in natural language. It’s possible that at this stage in language acquisition, learners still require a large degree of flexibility in their representations of grammatical rules. Furthermore, while our stimuli may not have a large degree of external validity, Experiments 4 and 5 demonstrate a clear effect of sleep on abstraction of grammatical patterns. Alongside Gomez et al. (2006), these studies also provide evidence that there is a shift in the effect of sleep on stabilization of grammatical patterns, with greater stabilization later in development.

Chapter 7: General Discussion

7.1 Encoding and Retention of Syllable Order

Regarding encoding of word forms during statistical learning, we found that infants are able to encode the order of syllables within words, discriminating words with the correct syllable order from words whose syllables switched positions. Outside of temporal order, researchers have proposed that younger infants acquire word forms with a great level of detail. For example, 7.5-month-old infants who listened to the target word “cup” repeated by a female voice only recognize the same word when it was presented in the same voice. Infants did not recognize “cup” when a male voice produced the word. By 10 months, infants form more abstract, flexible representations of word forms and can recognize the target word when it is presented in different speaker voices (Houston & Jusczyk, 2000). Though these studies specifically investigate infants’ perceptual specificity, asking whether they will discriminate based on talker voice rather than changes in the word form, it follows that infants would also encode details more important for future language use such as the order of syllables within words.

Our results with adult participants are less clear. We found a clear effect of the time of consolidation on learning. Participants who were trained in the morning showed poor performance on immediate test, endorsing many illegal items (PWs and BWs) and FWs. Contrastingly, participants who listened to the artificial language in the evening endorsed forwards words at a greater rate than both part-words and backwards-order words, showing they were sensitive to both co-occurrence frequency and order at immediate test. This pattern of results indicates participants show increased statistical learning in the evening, before a period of sleep.

Considering retention of order, our preliminary findings with 13-month-olds infants indicate sleeping after learning does not lead to retention of order. In this study, the 8 infants who napped after learning did not discriminate words with correct syllable order from words with incorrect syllable order. Similarly, adults who slept after learning did not show retention of order at the delayed test at the group level. These findings suggest sleep does not benefit memory for order after statistical learning in infancy or adulthood, going against research on declarative memory for the order of word associations (Drosopoulos et al., 2007). However, our data with adults revealed a correlation between total sleep time and retention of order at the delayed test, suggesting participants who slept more showed greater retention of order than those who did not. Here we provide evidence that sleeping influences memory for the order of syllables within words after word segmentation. Despite this correlation, as a group the participants who slept after learning could not reliably endorse forwards words more often than the illegal items, indicating poor retention overall with sleep. Furthermore, correlation analyses indicated participants in the wakefulness groups retained what they encoded during familiarization. This pattern of results indicates wakefulness participants retained while sleep participants as a group did not, though it is possible our sleep participants would have performed better if a greater number had slept longer between learning and test. Furthermore, it appears that sleep confers a benefit to retention of order over and above encoding strength, as immediate performance did not correlate with retention in the sleep group.

7.2 Veridical Memory for Nonadjacent Dependencies After Sleep

Focusing on retention of nonadjacent dependencies, 18-month-old infants benefit from napping soon after learning. In Experiment 4, infants who napped remember the exact NADs heard 4 hours earlier during familiarization. They can discriminate these legal sequences from phrases that violate the NADs heard before their nap. Gomez and colleagues (2006) reported a different pattern of results with 15-month-old infants, who showed a generalization effect after napping. Thus, it appears that sleep does not promote specific representations of the learned NADs at 15 months (Gomez et al., 2006), but it does at 18 months. Our preliminary findings with 13-month-olds are somewhat consistent with this finding, as these younger infants did not remember the order of syllables within words across sleep (though we will continue data collection to reassess retention of order after a nap). Taken together, our results suggest that in development (at 13 and 15 months), sleep does not promote the consolidation of detailed linguistic representations. By 18 months, sleep confers a benefit to veridical representations of nonadjacent dependencies. However, we are not yet certain whether sleep will benefit retention of syllable order in older infants. While we have not assessed when in development sleep benefits retention of temporal order in language, it is clear that the role of sleep in consolidation of NADs changes between 15 and 18 months of age.

7.3 Generalization After Sleep

As stated, 18-month-olds show veridical retention of NADs across sleep. What about generalization, which is an important aspect of language acquisition? A speaker would have great difficulty if they could not generalize their knowledge of grammar to instances where some of the vocabulary is novel. We followed up on this question by presenting NADs with completely novel vocabulary at test during Experiment 5. When infants listen to a language with a *guf-Y-zam* and a *miv-Y-fop* dependency during familiarization, 4 hours later they can recognize the NAD pattern in phrases with new lexical items, such as *pel-X-jic* and *vot-X-rud* only if they sleep soon after learning. Fifteen-month-old infants who napped after learning NADs continued to show generalization 24 hours later when tested with the same vocabulary (Hupbach et al., 2009). Following this line of thought, we predict that 18-month-olds would continue to show generalization to new vocabulary after 24 hours, theorizing that sleep promotes long-lasting abstract representations for grammatical patterns.

In Experiments 3, 4, and 5, we utilized difference scores to assess retention, calculating the average response to consistent or familiar test stimuli and subtracting average responses to unfamiliar or inconsistent test stimuli. ERP and fMRI research commonly use subtraction in order to isolate a desired effect. For example, calculating the difference between ERPs elicited from semantically correct versus semantically incorrect sentences can isolate the signature of the "semantic error" response. In our studies, the difference value does not have such a straightforward meaning. In Experiment 3, to calculate correlations between immediate and delayed performance, we subtracted endorsement rates of backwards-order and part-words from endorsement rates of forwards words. This provided us with a measurement reflecting the degree to which participants endorsed legal items to a greater extent than illegal items, an indication of learning. Similarly, when calculating difference scores related to looking time in our studies with 18-month-olds, our measure reflects the difference in attention paid to one trial type versus the other. In contrast to physiological research, we only implemented difference scores to concisely analyze our data rather than to reveal psychological constructs.

7.4 Changes in Sleep-dependent Consolidation with Development

One possible explanation for the observed changes in memory retention with development is the 18-month-olds' experience with NADs and other grammatical dependencies in natural language. At 18 months, infants can discriminate ungrammatical phrases like *can running* from grammatical phrases like *is running* (Santelmann & Jusczyk, 1998). Perhaps their greater knowledge of NADs in general facilitates sleep-dependent memory consolidation to some extent. This follows the findings from Horvath and colleagues (2015), who found infants with larger vocabularies retained more word forms across sleep than infants with smaller vocabularies. Furthermore, studies with adults and older children also support the idea that pre-existing knowledge confers a benefit to sleep-dependent memory consolidation (reviewed in [James et al., 2017]). It could be the case that having pre-existing knowledge that is related to the to-be-learned information may promote consolidation over a period of sleep. As infants at 18 months have more advanced NAD knowledge than infants at 15 months, this difference could yield more stabilization with sleep at 18 months.

The advances we observe from 15 months to 18 months could also be a result of encoding strength. Gómez (2002) illustrated 18-month-olds show a novelty effect immediately after listening to an artificial language containing NADs. As mentioned in our discussion of Experiment 1, researchers theorize that novelty effects reflect the infants have formed a robust representation of the familiarization stimulus 9/16/20 11:33 AM. On the other hand, 15-month-old infants demonstrate a familiarity effect immediately after learning NADs, a sign of less-mature encoding (Gómez & Maye, 2005). Along similar lines, 18-month-olds have a larger processing window than 15-month-olds, allowing older infants to hold more information in working memory. Though investigators have not empirically demonstrated working memory increases between these specific ages, studies show processing windows grow in the span of months during development (Reynolds & Romano, 2016, for review). Furthermore, Santelmann and Jusczyk (1998) found that 18-month-olds cannot recognize NADs in English when the two related elements are separated by five syllables, but they can when the related elements are separated by three syllables. This study provides evidence that infants' recognition of NADs is limited by the size of their processing window, especially when many interleaving elements come between the associated items. Increases in working memory with age might allow older infants to represent an entire NAD in memory at once. Shorter working memory span, on the other hand, might allow an infant to hold only part of the NAD in mind, preventing them from forming robust representations (Santelmann & Jusczyk, 1998). Therefore, age-related differences in acquisition of NADs could be responsible for the increased benefits of sleep with development, especially considering word learning studies show strongly encoded memories are consolidated to a greater degree than weakly encoded memories (James et al., 2017; Walker et al., 2019).

Alternatively, the neural systems that drive sleep-dependent memory consolidation may be not be mature enough by 15 months to benefit veridical retention over sleep; it is possible that neural development and changes in sleep physiology between 15 and 18 months enhance the effect of sleep on memory. For example, Simon et al. (2017) found that 6.5-month-olds infants who showed the most mature sleep physiology during a post-learning nap demonstrated more robust retention over a delay of 90 minutes. However, it is important to note that the mechanisms behind memory consolidation during sleep, especially those functioning during early development, are not certain.

Generally, research with adult participants shows memories are strengthened over sleep, especially if they are encoded with the involvement of the hippocampus (Marshall & Born, 2007). Moreover, many studies support the notion that activity within the hippocampus during sleep supports memory consolidation. Namely, studies with mice revealed the phenomenon referred to as "neural replay," which occurs when the neurons that were active during learning repeat the same activations during subsequent sleep (Pavlides & Winson, 1989; Skaggs & McNaughton, 1996). Neural replay, along with the synchronous activity between the hippocampus and neocortex, assists strengthening of information in cortical memory. This theory of sleep-dependent memory consolidation is called Active Systems Consolidation (reviewed in Born & Wilhelm, 2012). Furthermore, Lewis & Durrant (2011) propose neural replay has a role in generalization: after repeated activations of learned representations, the shared, consistent aspects of these memories are strengthened. In this way, neural replay supports integration across multiple exemplars.

Critically, the hippocampal structures driving neural replay are slow to develop relative to other brain regions that support memory, such as cortical systems. Considering the structures supporting sleep-dependent memory consolidation take years to develop (Lavenex & Banta Lavenex, 2013), it is possible that young infants do not have sufficient connectivity to benefit from memory stabilization across sleep until their neural systems have reached a certain state in maturation (Gomez & Edgin, 2015). According to Active Systems Consolidation, 18-month-olds observe greater benefit from sleep as they have more mature connectivity both within the hippocampus and between the hippocampus and cortical structures than 15-month-olds. Advanced connectivity relative to 15-month-olds allows the older infants to consolidate the exact vocabulary heard in the NADs during familiarization. At the same time, repeated activation of the NAD representations facilitates integration across the exemplars, so that the consistent aspect of all NADs from familiarization (the relationship between the 1st and 3rd word) is retained and strengthened with sleep. Fifteen-month-olds, on the other hand, do not benefit from hippocampal processing during sleep, so they forget the specific vocabulary making up the NADs. This targeted forgetting of the vocabulary paired with retention of the NAD relationship could be a result of synaptic homeostasis, which proposes recently formed connections are weakened during sleep (Tononi & Cirelli, 2006). We will return to this idea later on in this discussion.

7.4.1 The Hippocampus, Statistical Learning, and Sleep

Along these lines, it is not clear whether statistical word segmentation (or statistical learning in general) is supported by the hippocampus, though a number of studies assessing statistical learning and word segmentation in people with hippocampal amnesia provide evidence that the hippocampus is involved in this process insofar as patients performed at chance or at significantly lower levels than healthy controls (Covington, Brown-Schmidt, & Duff, 2018; Schapiro, Gregory, Landau, McCloskey & Turke-Browne, 2014). These studies indicate the hippocampus is active during statistical learning, but we did not find sleep to benefit retention of co-occurrence frequency in Experiment 3 (though we only assess participants behaviorally, so we cannot rule out hippocampal involvement). Similarly, research using fMRI revealed cortical areas such as the inferior frontal and temporal lobe track statistical information, alongside subcortical structures such as the caudate rather than the hippocampus (Karuza et al., 2013; McNealy et al., 2006; Plante et al., 2015).

If the hippocampus does not play a vital role during statistical word learning, hippocampal-dependent processes during sleep may not function to strengthen these memories the same way they consolidate explicit memories, which rely on the hippocampus to a greater degree. Our study with adults suggests sleep after learning is linked to retention of syllable order, rather than co-occurrence frequency, indicating sleep specifically benefits memory for the order of learned information above sequential statistics. Along these lines, imaging work with visual statistical learning tasks suggest the hippocampus preferentially encodes forwards transitions during statistical segmentation but not backwards transitions (Schapiro et al., 2012, 2017). This is also line with studies on explicit learning, in which sleep strengthens memory for forwards transitions but not backwards transitions (Drosopoulos et al., 2007). It could be the case that the hippocampus is active during statistical learning because it is tracking the ordering of items, and this information is then consolidated with later sleep.

Furthermore, it is possible that other mechanisms, independent of the hippocampus, function to consolidate statistical learning across sleep. Our findings with 18-month-olds follows this somewhat, as these infants benefit from sleep while still having a relatively immature hippocampus (Lavenex & Banta Lavenex, 2013). One possible mechanism for sleep-dependent consolidation in infancy and cortical memories is synaptic homeostasis. This theory proposes that active synaptic connections become strengthened during experiences across wakefulness. The connections that are used most build up the most "weight". Then, during sleep, there is a global downscaling in strength, so that only the strongest connections remain after sleep. This type of consolidation is thought to save energy and promote plasticity by preserving neural "space" for the most important synaptic connections (Tononi & Cirelli, 2006).

Following the synaptic homeostasis theory, we might see an improvement in memory stabilization in 18-month-olds relative to 15-month-olds because older infants form stronger connections during learning, which persist across downscaling over sleep. If infants form two levels of representation, they may encode both the positional information (Benavides-Varela & Mehler, 2015), like the specific vocabulary that occurs at the beginning and end of phrases, along with the general NAD rule that the first and third items are associated. However, 18-month-olds strongly encode both of these aspects, while 15-month-olds' representation of the general NAD relationship is stronger than that of the specific vocabulary. This is because the NAD relationship was consistent across the entire familiarization grammar, but the specific vocabulary making up the NADs varied. Then, across sleep, fifteen-month-olds' knowledge of the vocabulary is pruned because these connections were weaker and did not survive downscaling, leaving only the NAD relationship. Thus, these infants generalize to NADs that violate those from familiarization because they do not retain the specific vocabulary (Gomez et al., 2006). At 18 months, infants form stronger connections during learning, so both levels of representation persist across synaptic downscaling during sleep. Thus, after napping, 18-month-olds are able to recognize the NADs they heard during familiarization (Experiment 4). At the same time, they retain the general association between the first and third words in each phrase, not specific to the vocabulary heard during familiarization, allowing them to generalize to new vocabulary (Experiment 5).

7.5 Concluding Remarks

Overall, we provide evidence that sleep supports consolidation of information gained through statistical learning early in development. Namely, we found that sleep promotes specific memory and abstraction of a grammatical rule in 18-month-olds, though the mechanism behind generalization is currently unclear. With continued data collection, we believe infants who stay awake after learning will retain specific representations, but our current data strongly suggest they cannot generalize across wakefulness. Infants who sleep, on the other hand, generalize NADs to new vocabulary. Thus, sleep facilitates abstraction in 18-month-old infants. Furthermore, comparing our results to those with 15-month-olds (Gomez et al, 2006), we find a shift in the role of sleep: earlier in development, sleep benefits abstraction but not specific memory. By 18 months, sleep benefits both abstraction and specific memory.

Furthermore, we found sleep relates to retention of syllable order after word segmentation in adulthood. However, sleep did not relate to retention of co-occurrence frequency, and participants who slept after learning did not demonstrate retention as a group (similar to our preliminary findings with 13-month-old infants). On the other hand, adults who remained awake after learning successfully retained what they encoded before the delay. This pattern of results suggests sleep may promote retention of order in adulthood, though memory for statistics does not appear to benefit from sleep. Perhaps statistical information is consolidated through processes like synaptic homeostasis, so only the strongest connections formed during learning survive across a period of sleep.

Future studies using polysomnography in adults can provide further information on the aspects of sleep that support consolidation of order after statistical learning. For example, slow-wave activity is an indicator of synaptic downscaling, and slow-wave activity increases during sleep after learning (Tononi & Cirelli, 2006), with specific local increases in regions that had recently been active during learning (Huber et al., 2004). Therefore, if processes related to synaptic downscaling acted to consolidate statistical learning, we may observe increased SWA over temporal and potentially frontal regions, as these have been implicated in statistical language learning (Plante et al., 2015).

In addition, further testing may reveal what drives the changes in the observed effects of sleep between 15 and 18 months of age. For example, designs that measure strength of encoding can elucidate whether older infants show greater sleep-dependent consolidation due to more advanced encoding during learning. Furthermore, polysomnography studies would allow us to measure whether there are significant differences in sleep physiology between 15 and 18 months, as well as correlate aspects of physiology to retention and generalization. Existing studies indicate slow-wave activity increases in the first years of life and declines after puberty (Feinberg et al., 1977), and the topographic distribution of slow-wave activity changes with age, moving from posterior regions to frontal regions with maturation (Kurth et al., 2010). In addition, greater slow-wave activity in frontal regions is associated with more robust consolidation across sleep in 6.5-month-old infants (Simon et al., 2017). Finally, children with a higher ratio of frontal to posterior slow-wave activity have greater myelin globally, suggesting more mature neurobiology. This study indicates the ratio of frontal to posterior slow-wave activity can be used to infer maturational state (LeBourgeois et al., 2019). Considering sleep physiology undergoes drastic changes during development, more mature physiology could support consolidation to a greater extent. More specifically, we might expect more 18-month-old infants to show increased power in their frontal slow-wave activity or in their ratio of frontal to posterior slow-wave activity in comparison to 15-month-olds, leading to more advanced consolidation. Comparing sleep physiology between these ages would allow us to pinpoint whether changes in specific processes during sleep, such as frontal slow-wave activity, promote both stabilization and abstraction as infants develop.

Appendix – Questionnaires

1. Morning-Eveningness Questionnaire *to determine participants' time-of-day preference.*

MORNINGNESS-EVENINGNESS QUESTIONNAIRE
Self-Assessment Version (MEQ-SA)[1]

Name: ______________________________ Date: _________________________

For each question, please select the answer that best describes you by circling the point value that best indicates how you have felt in recent weeks.

1. *Approximately* what time would you get up if you were entirely free to plan your day?

   [5] 5:00 AM–6:30 AM *(05:00–06:30 h)*
   [4] 6:30 AM–7:45 AM *(06:30–07:45 h)*
   [3] 7:45 AM–9:45 AM *(07:45–09:45 h)*
   [2] 9:45 AM–11:00 AM *(09:45–11:00 h)*
   [1] 11:00 AM–12 noon *(11:00–12:00 h)*

2. *Approximately* what time would you go to bed if you were entirely free to plan your evening?

   [5] 8:00 PM–9:00 PM *(20:00–21:00 h)*
   [4] 9:00 PM–10:15 PM *(21:00–22:15 h)*
   [3] 10:15 PM–12:30 AM *(22:15–00:30 h)*
   [2] 12:30 AM–1:45 AM *(00:30–01:45 h)*
   [1] 1:45 AM–3:00 AM *(01:45–03:00 h)*

3. If you usually have to get up at a specific time in the morning, how much do you depend on an alarm clock?

   [4] Not at all
   [3] Slightly
   [2] Somewhat
   [1] Very much

[1]Some stem questions and item choices have been rephrased from the original instrument (Horne and Östberg, 1976) to conform with spoken American English. Discrete item choices have been substituted for continuous graphic scales. Prepared by Terman M, Rifkin JB, Jacobs J, White TM (2001), New York State Psychiatric Institute, 1051 Riverside Drive, Unit 50, New York, NY, 10032. January 2008 version. Supported by NIH Grant MH42931. *See also:* automated version (AutoMEQ) at www.cet.org.

Horne JA and Östberg O. A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. International Journal of Chronobiology, 1976: 4, 97-100.

MORNINGNESS-EVENINGNESS QUESTIONNAIRE
Page 2

4. How easy do you find it to get up in the morning (when you are not awakened unexpectedly)?

   [1] Very difficult
   [2] Somewhat difficult
   [3] Fairly easy
   [4] Very easy

5. How alert do you feel during the first half hour after you wake up in the morning?

   [1] Not at all alert
   [2] Slightly alert
   [3] Fairly alert
   [4] Very alert

6. How hungry do you feel during the first half hour after you wake up?

   [1] Not at all hungry
   [2] Slightly hungry
   [3] Fairly hungry
   [4] Very hungry

7. During the first half hour after you wake up in the morning, how do you feel?

   [1] Very tired
   [2] Fairly tired
   [3] Fairly refreshed
   [4] Very refreshed

8. If you had no commitments the next day, what time would you go to bed compared to your usual bedtime?

   [4] Seldom or never later
   [3] Less that 1 hour later
   [2] 1-2 hours later
   [1] More than 2 hours later

MORNINGNESS-EVENINGNESS QUESTIONNAIRE
Page 3

9. You have decided to do physical exercise. A friend suggests that you do this for one hour twice a week, and the best time for him is between 7-8 AM *(07-08 h)*. Bearing in mind nothing but your own internal "clock," how do you think you would perform?

   [4] Would be in good form
   [3] Would be in reasonable form
   [2] Would find it difficult
   [1] Would find it very difficult

10. At *approximately* what time in the evening do you feel tired, and, as a result, in need of sleep?

   [5] 8:00 PM–9:00 PM *(20:00–21:00 h)*
   [4] 9:00 PM–10:15 PM *(21:00–22:15 h)*
   [3] 10:15 PM–12:45 AM *(22:15–00:45 h)*
   [2] 12:45 AM–2:00 AM *(00:45–02:00 h)*
   [1] 2:00 AM–3:00 AM *(02:00–03:00 h)*

11. You want to be at your peak performance for a test that you know is going to be mentally exhausting and will last two hours. You are entirely free to plan your day. Considering only your "internal clock," which one of the four testing times would you choose?

   [6] 8 AM–10 AM *(08–10 h)*
   [4] 11 AM–1 PM *(11–13 h)*
   [2] 3 PM–5 PM *(15–17 h)*
   [0] 7 PM–9 PM *(19–21 h)*

12. If you got into bed at 11 PM *(23 h),* how tired would you be?

   [0] Not at all tired
   [2] A little tired
   [3] Fairly tired
   [5] Very tired

MORNINGNESS-EVENINGNESS QUESTIONNAIRE
Page 4

13. For some reason you have gone to bed several hours later than usual, but there is no need to get up at any particular time the next morning. Which one of the following are you most likely to do?

    [4] Will wake up at usual time, but will not fall back asleep
    [3] Will wake up at usual time and will doze thereafter
    [2] Will wake up at usual time, but will fall asleep again
    [1] Will not wake up until later than usual

14. One night you have to remain awake between 4-6 AM *(04-06 h)* in order to carry out a night watch. You have no time commitments the next day. Which one of the alternatives would suit you best?

    [1] Would not go to bed until the watch is over
    [2] Would take a nap before and sleep after
    [3] Would take a good sleep before and nap after
    [4] Would sleep only before the watch

15. You have two hours of hard physical work. You are entirely free to plan your day. Considering only your internal "clock," which of the following times would you choose?

    [4] 8 AM–10 AM *(08–10 h)*
    [3] 11 AM–1 PM *(11–13 h)*
    [2] 3 PM–5 PM *(15–17 h)*
    [1] 7 PM–9 PM *(19–21 h)*

16. You have decided to do physical exercise. A friend suggests that you do this for one hour twice a week. The best time for her is between 10-11 PM *(22-23 h)*. Bearing in mind only your internal "clock," how well do you think you would perform?

    [1] Would be in good form
    [2] Would be in reasonable form
    [3] Would find it difficult
    [4] Would find it very difficult

MORNINGNESS-EVENINGNESS QUESTIONNAIRE
Page 5

17. Suppose you can choose your own work hours. Assume that you work a five-hour day (including breaks), your job is interesting, and you are paid based on your performance. At *approximately* what time would you choose to begin?

    [5] 5 hours starting between 4–8 AM *(05–08 h)*
    [4] 5 hours starting between 8–9 AM *(08–09 h)*
    [3] 5 hours starting between 9 AM–2 PM *(09–14 h)*
    [2] 5 hours starting between 2–5 PM *(14–17 h)*
    [1] 5 hours starting between 5 PM–4 AM *(17–04 h)*

18. At *approximately* what time of day do you usually feel your best?

    [5] 5–8 AM *(05–08 h)*
    [4] 8–10 AM *(08–10 h)*
    [3] 10 AM–5 PM *(10–17 h)*
    [2] 5–10 PM *(17–22 h)*
    [1] 10 PM–5 AM *(22–05 h)*

19. One hears about "morning types" and "evening types." Which one of these types do you consider yourself to be?

    [6] Definitely a morning type
    [4] Rather more a morning type than an evening type
    [2] Rather more an evening type than a morning type
    [1] Definitely an evening type

____ Total points for all 19 questions

MORNINGNESS-EVENINGNESS QUESTIONNAIRE
Page 6

INTERPRETING AND USING YOUR MORNINGNESS-EVENINGNESS SCORE

This questionnaire has 19 questions, each with a number of points. First, add up the points you circled and enter your total morningness-eveningness score here:

Scores can range from 16-86. Scores of 41 and below indicate "evening types." Scores of 59 and above indicate "morning types." Scores between 42-58 indicate "intermediate types."

16-30	31-41	42-58	59-69	70-86
definite evening	moderate evening	intermediate	moderate morning	definite morning

Occasionally a person has trouble with the questionnaire. For example, some of the questions are difficult to answer if you have been on a shift work schedule, if you don't work, or if your bedtime is unusually late. Your answers may be influenced by an illness or medications you may be taking. *If you are not confident about your answers, you should also not be confident about the advice that follows.*

One way to check this is to ask whether your morningness-eveningness score approximately matches the sleep onset and wake-up times listed below:

Score	16-30	31-41	42-58	59-69	70-86
Sleep onset	2:00-3:00 AM *(02:00-03:00 h)*	12:45-2:00 AM *(00:45-02:00 h)*	10:45 PM-12:45 AM *(22:45-00:45 h)*	9:30-10:45 PM *(21:30-22:45 h)*	9:00-9:30 PM *(21:00-21:30 h)*
Wake-up	10:00-11:30 AM *(10:00-11:30 h)*	8:30-10:00 AM *(08:30-10:00 h)*	6:30-8:30 AM *(06:30-08:30 h)*	5:00-6:30 AM *(05:00-06:30 h)*	4:00-5:00 AM *(04:00-05:00 h)*

If your usual sleep onset is earlier than 9:00 PM *(21:00 h)* or later than 3:00 AM *(03:00 h)*, or your wake-up is earlier than 4:00 AM *(04:00 h)* or later than 11:30 AM *(11:30 h)*, you should seek the advice of a light therapy clinician in order to proceed effectively with treatment.

We use the morningness-eveningness score to improve the antidepressant effect of light therapy. Although most people experience good antidepressant response to light therapy when they take a regular morning session using a 10,000 lux white light device *(see www.cet.org for recommendations)* for 30 minutes, often this will not give the best possible response. If your internal clock is shifted relative to external time (as indirectly measured by your morningness-eveningness score), the timing of light therapy needs to be adjusted.

The table at the top of the next page shows the recommended start time for light therapy for a wide range of morningness-eveningness scores. If your score falls beyond this range (either very low or very high), you should seek the advice of a light therapy clinician in order to proceed effectively with treatment.

2. General Questionnaire

Mechanisms in Language- Post-experiment Questionnaire

1. Did any of the nonsense words stand to you out more than others? If so, which ones?

2. Can you recall any of the nonsense words?

3. On a scale from 1 to 5, how much attention did you pay to the language you were hearing?

	1	2	3	4	5
	not very much				very much

4. On a scale from 1 to 5, how hard were you trying when you gave judgments about whether strings were grammatical or not? (Please provide ratings for both part 1 and part 2 if applicable.)

Part 1: (if applicable)	1	2	3	4	5
	not very hard				very hard
Part 2:	1	2	3	4	5
	not very hard				very hard

5. On a scale from 1 to 5, how well did you feel during the experiment? (Please provide ratings for both part 1 and part 2.)

Part 1:	1	2	3	4	5
	not very well				very well
Part 2:	1	2	3	4	5
	not very well				very well

6. On a scale from 1 to 5, how alert did you feel during the experiment? (Please provide ratings for both part 1 and part 2.)

Part 1:	1	2	3	4	5
	not very alert				very alert
Part 2:	1	2	3	4	5
	not very alert				very alert

7. Have you had any caffeine in the last 8 hours or in the 8 hours before your participation in part 1 (e.g. coffee, tea, caffeinated soda, chocolate bar)? If so, what and when?

8. Have you had any alcohol in the last 24 hours? (Y/N)

9. Have you had any medications that might influence your alertness? (Y/N)

10. What did you think the experiment was about? Did you have any thoughts (general impressions) about the experiment while you were doing it?

11. Did you push the wrong button in response to any of the test items (during both part 1 and part 2 if applicable)? If so, about how many times did this happen?

12. Do you have any hearing loss?

13. Are you exposed to any language besides English? If so, how often?

14. Was your sleep last night normal compared to how you regularly sleep? If not, how was it different?

3. Sleep Log

Adult Sleep Log

Time	Mark Sleep/Wake times	Comments
9:00 p.m.		
9:30 p.m.		
10:00 p.m.		
10:30 p.m.		
11:00 p.m.		
11:30 p.m.		
12:00 a.m.		
12:30 p.m.		
1:00 a.m.		
1:30 a.m.		
2:00 a.m.		
2:30 a.m.		
3:00 a.m.		
3:30 a.m.		
4:00 a.m.		
4:30 a.m.		
5:00 a.m.		
5:30 a.m.		
6:00 a.m.		
6:30 a.m.		
7:00 a.m.		
7:30 a.m.		
8:00 a.m.		
8:30 a.m.		
9:00 a.m.		
9:30 a.m.		
10:00 a.m.		

Please mark next to the time you fell asleep last night and the time you woke up this morning.

References

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of Conditional Probability Statistics by 8-Month-Old Infants. *Psychological Science*, *9*(4), 321–324. https://doi.org/10.1111/1467-9280.00063

Aton, S. J., Seibt, J., Dumoulin, M., Jha, S. K., Steinmetz, N., Coleman, T., Naidoo, N., & Frank, M. G. (2009). Mechanisms of Sleep-Dependent Consolidation of Cortical Plasticity. *Neuron*, *61*(3), 454–466. https://doi.org/10.1016/j.neuron.2009.01.007

Batterink, L. J., Oudiette, D., Reber, P. J., & Paller, K. A. (2014). Sleep facilitates learning a new linguistic rule. *Neuropsychologia*, *65*, 169–179. https://doi.org/10.1016/j.neuropsychologia.2014.10.024

Batterink, L. J., & Paller, K. A. (2017). Sleep-Based Memory Processing Facilitates Grammatical Generalization: Evidence from Targeted Memory Reactivation. *Brain and Language*, *167*, 83–93. https://doi.org/10.1016/j.bandl.2015.09.003

Benavides‐Varela, S., & Mehler, J. (2015). Verbal Positional Memory in 7-Month-Olds. *Child Development*, *86*(1), 209–223. https://doi.org/10.1111/cdev.12291

Berger, S. E., & Scher, A. (2017). Naps improve new walkers' locomotor problem solving. *Journal of Experimental Child Psychology*, *162*, 292–300. https://doi.org/10.1016/j.jecp.2017.04.021

Born, J., & Wilhelm, I. (2012). System consolidation of memory during sleep. *Psychological Research*, *76*(2), 192–203. https://doi.org/10.1007/s00426-011-0335-6

Courage, M. L., & Howe, M. L. (2001). Long-Term Retention in 3.5-Month-Olds: Familiarization Time and Individual Differences in Attentional Style. *Journal of Experimental Child Psychology*, *79*(3), 271–293. https://doi.org/10.1006/jecp.2000.2606 de Hevia, M. D., & Spelke, E. S. (2010). Number-Space Mapping in Human Infants. *Psychological Science*, *21*(5), 653–660. https://doi.org/10.1177/0956797610366091

Delpouve, J., Schmitz, R., & Peigneux, P. (2014). Implicit learning is better at subjectively defined non-optimal time of day. *Cortex*, *58*, 18–22. https://doi.org/10.1016/j.cortex.2014.05.006

Diekelmann, S., & Born, J. (2010). The memory function of sleep. *Nature Reviews Neuroscience*, *11*(2), 114–126. https://doi.org/10.1038/nrn2762

Doyon, J., Korman, M., Morin, A., Dostie, V., Tahar, A. H., Benali, H., Karni, A., Ungerleider, L. G., & Carrier, J. (2009). Contribution of night and day sleep vs. Simple passage of time to the consolidation of motor sequence and visuomotor adaptation learning. *Experimental Brain Research*, *195*(1), 15–26. https://doi.org/10.1007/s00221-009-1748-y

Drosopoulos, S., Windau, E., Wagner, U., & Born, J. (2007). Sleep Enforces the Temporal Order in Memory. *PLOS ONE*, *2*(4), e376. https://doi.org/10.1371/journal.pone.0000376

Dumay, N., & Gaskell, M. G. (2007). Sleep-Associated Changes in the Mental Representation of Spoken Words. *Psychological Science*, *18*(1), 35–39. https://doi.org/10.1111/j.1467-9280.2007.01845.x

Feinberg, I., Hibi, S., & Carlson, V. R. (1977). Changes in EEG Amplitude During Sleep with Age. In K. Nandy & I. Sherwin (Eds.), *The Aging Brain and Senile Dementia* (Vol. 23, pp. 85–98). Springer US. https://doi.org/10.1007/978-1-4684-3093-6_6

Frank, M. C., Tenenbaum, J. B., & Gibson, E. (2013). Learning and Long-Term Retention of Large-Scale Artificial Languages. *PLOS ONE*, *8*(1), e52500. https://doi.org/10.1371/journal.pone.0052500

Frank, M. G., Issa, N. P., & Stryker, M. P. (2001). Sleep Enhances Plasticity in the Developing Visual Cortex. *Neuron*, *30*(1), 275–287. https://doi.org/10.1016/S0896-6273(01)00279-3

Friederici, A. D., Mueller, J. L., Sehm, B., & Ragert, P. (2013). Language Learning without Control: The Role of the PFC. *Journal of Cognitive Neuroscience*, *25*(5), 814–821. https://doi.org/10.1162/jocn_a_00350

Friedrich, M., Wilhelm, I., Born, J., & Friederici, A. D. (2015). Generalization of word meanings during infant sleep. *Nature Communications*, *6*(1), 6004. https://doi.org/10.1038/ncomms7004

Gómez, R. L. (2002). Variability and Detection of Invariant Structure. *Psychological Science*, *13*(5), 431–436. https://doi.org/10.1111/1467-9280.00476

Gómez, R. L., Bootzin, R. R., & Nadel, L. (2006). Naps Promote Abstraction in Language-Learning Infants. *Psychological Science*, *17*(8), 670–674. https://doi.org/10.1111/j.1467-9280.2006.01764.x

Gómez, R., & Maye, J. (2005). The Developmental Trajectory of Nonadjacent Dependency Learning. *Infancy*, *7*(2), 183–206. https://doi.org/10.1207/s15327078in0702_4

Graf Estes, K., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can Infants Map Meaning to Newly Segmented Words?: Statistical Segmentation and Word Learning. *Psychological Science*, *18*(3), 254–260. https://doi.org/10.1111/j.1467-9280.2007.01885.x

Hay, J. F., Pelucchi, B., Graf Estes, K., & Saffran, J. R. (2011). Linking sounds to meanings: Infant statistical learning in a natural language. *Cognitive Psychology*, *63*(2), 93–106. https://doi.org/10.1016/j.cogpsych.2011.06.002

Henderson, L. M., Weighall, A. R., Brown, H., & Gaskell, M. G. (2012). Consolidation of vocabulary is associated with sleep in children. *Developmental Science*, *15*(5), 674–687. https://doi.org/10.1111/j.1467-7687.2012.01172.x

Horne, J. A., & Östberg, O. (1976). A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *International Journal of Chronobiology*, *4*, 97–110.

Horváth, K., Hannon, B., Ujma, P. P., Gombos, F., & Plunkett, K. (2018). Memory in 3-month-old infants benefits from a short nap. *Developmental Science*, *21*(3), e12587. https://doi.org/10.1111/desc.12587

Horváth, K., Myers, K., Foster, R., & Plunkett, K. (2015). Napping facilitates word learning in early lexical development. *Journal of Sleep Research*, *24*(5), 503–509. https://doi.org/10.1111/jsr.12306

Houston, D. M., & Jusczyk, P. W. (2000). The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(5), 1570–1582. https://doi.org/10.1037/0096-1523.26.5.1570

Houston-Price, C., & Nakai, S. (2004). Distinguishing novelty and familiarity effects in infant preference procedures. *Infant and Child Development*, *13*(4), 341–348. https://doi.org/10.1002/icd.364

Huber, R., Felice Ghilardi, M., Massimini, M., & Tononi, G. (2004). Local sleep and learning. *Nature*, *430*(6995), 78–81. https://doi.org/10.1038/nature02663

Hunter, M. A., Ames, E. W., & Koopman, R. (1983). Effects of stimulus complexity and familiarization time on infant preferences for novel and familiar stimuli. *Developmental Psychology*, *19*(3), 338–352. https://doi.org/10.1037/0012-1649.19.3.338

Hupbach, A., Gomez, R. L., Bootzin, R. R., & Nadel, L. (2009). Nap-dependent learning in infants. *Developmental Science*, *12*(6), 1007–1012. https://doi.org/10.1111/j.1467-7687.2009.00837.x

James, E., Gaskell, M. G., Weighall, A., & Henderson, L. (2017). Consolidation of vocabulary during sleep: The rich get richer? *Neuroscience & Biobehavioral Reviews*, *77*, 1–13. https://doi.org/10.1016/j.neubiorev.2017.01.054

Jha, S. K. (2005). Sleep-Dependent Plasticity Requires Cortical Activity. *Journal of Neuroscience*, *25*(40), 9266–9274. https://doi.org/10.1523/JNEUROSCI.2722-05.2005

Karuza, E. A., Newport, E. L., Aslin, R. N., Starling, S. J., Tivarus, M. E., & Bavelier, D. (2013). The neural correlates of statistical learning in a word segmentation task: An fMRI study. *Brain and Language*, *127*(1), 46–54. https://doi.org/10.1016/j.bandl.2012.11.007

Kemler Nelson, D. G., Jusczyk, P. W., Mandel, D. R., Myers, J., Turk, A., & Gerken, L. (1995). The head-turn preference procedure for testing auditory perception. *Infant Behavior and Development*, *18*(1), 111–116. https://doi.org/10.1016/0163-6383(95)90012-8

Kurczek, J., Brown-Schmidt, S., & Duff, M. C. (2013). Hippocampal contributions to language: Evidence of referential processing deficits in amnesia. *Journal of Experimental Psychology. General*, *142*(4), 1346–1354. https://doi.org/10.1037/a0034026

Kurth, S., Ringli, M., Geiger, A., LeBourgeois, M., Jenni, O. G., & Huber, R. (2010). Mapping of Cortical Activity in the First Two Decades of Life: A High-Density Sleep Electroencephalogram Study. *Journal of Neuroscience*, *30*(40), 13211–13219. https://doi.org/10.1523/JNEUROSCI.2532-10.2010

Lavenex, P., & Banta Lavenex, P. (2013). Building hippocampal circuits to learn and remember: Insights into the development of human memory. *Behavioural Brain Research*, *254*, 8–21. https://doi.org/10.1016/j.bbr.2013.02.007

LeBourgeois, M. K., Dean, D. C., Deoni, S. C. L., Kohler, M., & Kurth, S. (2019). A simple sleep EEG marker in childhood predicts brain myelin 3.5 years later. *NeuroImage*, *199*, 342–350. https://doi.org/10.1016/j.neuroimage.2019.05.072

Lewis, P. A., & Durrant, S. J. (2011). Overlapping memory replay during sleep builds cognitive schemata. *Trends in Cognitive Sciences*, *15*(8), 343–351. https://doi.org/10.1016/j.tics.2011.06.004

Lewkowicz, D. J. (2013). Development of Ordinal Sequence Perception in Infancy. *Developmental Science*, *16*(3), 352–364. https://doi.org/10.1111/desc.12029

Li, W., Ma, L., Yang, G., & Gan, W.-B. (2017). REM sleep selectively prunes and maintains new synapses in development and learning. *Nature Neuroscience*, *20*(3), 427–437. https://doi.org/10.1038/nn.4479

Louie, K., & Wilson, M. A. (2001). Temporally structured replay of awake hippocampal ensemble activity during rapid eye movement sleep. *Neuron*, *29*(1), 145–156. https://doi.org/10.1016/s0896-6273(01)00186-6

Marshall, L., & Born, J. (2007). The contribution of sleep to hippocampus-dependent memory consolidation. *Trends in Cognitive Sciences*, *11*(10), 442–450. https://doi.org/10.1016/j.tics.2007.09.001

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*(3), B101–B111. https://doi.org/10.1016/S0010-0277(01)00157-3

McNealy, K., Mazziotta, J. C., & Dapretto, M. (2006). Cracking the Language Code: Neural Mechanisms Underlying Speech Parsing. *Journal of Neuroscience*, *26*(29), 7629–7639. https://doi.org/10.1523/JNEUROSCI.5501-05.2006

Mirman, D., Magnuson, J. S., Estes, K. G., & Dixon, J. A. (2008). The link between statistical segmentation and word learning in adults. *Cognition*, *108*(1), 271–280. https://doi.org/10.1016/j.cognition.2008.02.003

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, *48*(2), 127–162. https://doi.org/10.1016/S0010-0285(03)00128-2

Nguyen, N. D., Tucker, M. A., Stickgold, R., & Wamsley, E. J. (2013). Overnight Sleep Enhances Hippocampus-Dependent Aspects of Spatial Memory. *Sleep*, *36*(7), 1051–1057. https://doi.org/10.5665/sleep.2808

Pavlides, C., & Winson, J. (1989). Influences of hippocampal place cell firing in the awake state on the activity of these cells during subsequent sleep episodes. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *9*(8), 2907–2918.

Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009). Learning in reverse: Eight-month-old infants track backward transitional probabilities. *Cognition*, *113*(2), 244–247. https://doi.org/10.1016/j.cognition.2009.07.011

Plante, E., Patterson, D., Gómez, R., Almryde, K. R., White, M. G., & Asbjørnsen, A. E. (2015). The nature of the language input affects brain activation during learning from a natural language. *Journal of Neurolinguistics*, *36*, 17–34. https://doi.org/10.1016/j.jneuroling.2015.04.005

Rasch, B., Buchel, C., Gais, S., & Born, J. (2007). Odor Cues During Slow-Wave Sleep Prompt Declarative Memory Consolidation. *Science*, *315*(5817), 1426–1429. https://doi.org/10.1126/science.1138581

Reynolds, G. D., & Romano, A. C. (2016). The Development of Attention Systems and Working Memory in Infancy. *Frontiers in Systems Neuroscience*, *10*. https://doi.org/10.3389/fnsys.2016.00015

Roediger, H. L., & Karpicke, J. D. (2006). Test-Enhanced Learning: Taking Memory Tests Improves Long-Term Retention. *Psychological Science*, *17*(3), 249–255. https://doi.org/10.1111/j.1467-9280.2006.01693.x

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical Learning by 8-Month-Old Infants. *Science*, *274*(5294), 1926–1928. https://doi.org/10.1126/science.274.5294.1926

Saffran, Jenny R., Newport, E. L., & Aslin, R. N. (1996). Word Segmentation: The Role of Distributional Cues. *Journal of Memory and Language*, *35*(4), 606–621. https://doi.org/10.1006/jmla.1996.0032

Santelmann, L. M., & Jusczyk, P. W. (1998). Sensitivity to discontinuous dependencies in language learners: Evidence for limitations in processing space. *Cognition*, *69*(2), 105–134. https://doi.org/10.1016/S0010-0277(98)00060-2

Schapiro, A. C., Kustner, L. V., & Turk-Browne, N. B. (2012). Shaping of Object Representations in the Human Medial Temporal Lobe Based on Temporal Regularities. *Current Biology*, *22*(17), 1622–1627. https://doi.org/10.1016/j.cub.2012.06.056

Schapiro, A. C., Turk-Browne, N. B., Botvinick, M. M., & Norman, K. A. (2017). Complementary learning systems within the hippocampus: A neural network modelling approach to reconciling episodic memory with statistical learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1711), 20160049. https://doi.org/10.1098/rstb.2016.0049

Seibt, J., Dumoulin, M. C., Aton, S. J., Coleman, T., Watson, A., Naidoo, N., & Frank, M. G. (2012). Protein Synthesis during Sleep Consolidates Cortical Plasticity In Vivo. *Current Biology*, *22*(8), 676–682. https://doi.org/10.1016/j.cub.2012.02.016

Simon, K. N. S., Werchan, D., Goldstein, M. R., Sweeney, L., Bootzin, R. R., Nadel, L., & Gómez, R. L. (2017). Sleep confers a benefit for retention of statistical language learning in 6.5month old infants. *Brain and Language*, *167*, 3–12. https://doi.org/10.1016/j.bandl.2016.05.002

Skaggs, W. E., & McNaughton, B. L. (1996). Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience. *Science (New York, N.Y.)*, *271*(5257), 1870–1873. https://doi.org/10.1126/science.271.5257.1870

Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*(3), 1558–1568. https://doi.org/10.1016/j.cognition.2007.06.010

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*(43), 14356–14360. https://doi.org/10.1523/JNEUROSCI.3028-10.2010

Thompson, S. P., & Newport, E. L. (2007). Statistical Learning of Syntax: The Role of Transitional Probability. *Language Learning and Development*, *3*(1), 1–42. https://doi.org/10.1080/15475440709336999

Tononi, G., & Cirelli, C. (2006). Sleep function and synaptic homeostasis. *Sleep Medicine Reviews*, *10*(1), 49–62. https://doi.org/10.1016/j.smrv.2005.05.002

Vlach, H. A., & Johnson, S. P. (2013). Memory constraints on infants' cross-situational statistical learning. *Cognition*, *127*(3), 375–382. https://doi.org/10.1016/j.cognition.2013.02.015

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*(1), 205–211. https://doi.org/10.1016/S0896-6273(02)00746-8

Walker, S., Henderson, L. M., Fletcher, F. E., Knowland, V. C. P., Cairney, S. A., & Gaskell, M. G. (2019). Learning to live with interfering neighbours: The influence of time of learning and level of encoding on word learning. *Royal Society Open Science*, *6*(4), 181842. https://doi.org/10.1098/rsos.181842

Wang, H.-C., Savage, G., Gaskell, M. G., Paulin, T., Robidoux, S., & Castles, A. (2017). Bedding down new words: Sleep promotes the emergence of lexical competition in visual word recognition. *Psychonomic Bulletin & Review*, *24*(4), 1186–1193. https://doi.org/10.3758/s13423-016-1182-7

Williams, S. E., & Horst, J. S. (2014). Goodnight book: Sleep consolidation improves word learning via storybooks. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00184