

Southern Illinois University Carbondale

OpenSIUC

---

Articles

Department of Linguistics

---

2012

## The roles of phonological short-term memory and working memory in L2 grammar and vocabulary learning

Katherine I. Martin  
martinki@siu.edu

Nick C. Ellis  
*University of Michigan - Ann Arbor*

Follow this and additional works at: [https://opensiuc.lib.siu.edu/ling\\_articles](https://opensiuc.lib.siu.edu/ling_articles)

---

### Recommended Citation

Martin, Katherine I. and Ellis, Nick C. "The roles of phonological short-term memory and working memory in L2 grammar and vocabulary learning." *Studies in Second Language Acquisition* 34, No. 3 (Jan 2012): 379-413. doi:10.1017/S0272263112000125.

This Article is brought to you for free and open access by the Department of Linguistics at OpenSIUC. It has been accepted for inclusion in Articles by an authorized administrator of OpenSIUC. For more information, please contact [opensiuc@lib.siu.edu](mailto:opensiuc@lib.siu.edu).

RUNNING HEAD: PHONOLOGICAL STM AND WORKING MEMORY IN L2  
LEARNING

The Roles of Phonological STM and Working Memory in L2 Grammar and Vocabulary  
Learning

Katherine I. Martin and Nick C. Ellis

The University of Michigan

## Abstract

This study analyzed phonological short-term memory (PSTM) and working memory (WM) and their relationship with vocabulary and grammar learning in an artificial foreign language. Nonword repetition, nonword recognition, and listening span were used as memory measures. Participants learned the singular forms of vocabulary for an artificial foreign language before being exposed to plural forms in sentence contexts. Participants were tested on their ability to induce the grammatical forms and to generalize the forms to novel utterances. Individual differences in final abilities in vocabulary and grammar correlated between 0.44 and 0.76, depending on the measure. Despite these strong associations, the results demonstrated significant independent effects of PSTM and WM upon L2 vocabulary learning and upon L2 grammar learning, some of which were mediated by vocabulary and some of which were direct effects.

## The Roles of Phonological STM and Working Memory in L2 Grammar and Vocabulary Learning

Adults are differentially successful from each other in their attempts to learn a second language (L2). Individual differences in many domains, including motivation (Fan, 2003; Sanaoui, 1995), age (Birdsong & Molis, 2001; Johnson & Newport, 1989), working memory (WM; Harrington & Sawyer, 1992; Sunderman & Kroll, 2009), and phonological short-term memory (PSTM; Gathercole & Baddeley, 1993; Service, 1992), have all been proposed as reasons for this differential success. This research has primarily focused on vocabulary learning, reading comprehension, and fluency development, rather than the learning of grammatical patterns. This study therefore investigated the roles of working memory and phonological short-term memory as they separately affect grammar and vocabulary learning in an artificial foreign language.

### *A Definition of Working Memory*

Working memory is the ability to mentally store and manipulate information relevant to a task (Baddeley, 1998, 2003). There are two broadly separate approaches in the literature, one British and one North American (Baddeley, 1998; Just & Carpenter, 1992; van den Noort, Bosch, & Hugdahl, 2006); however, the distinctions between them are not always made clear (Williams, 2012). In this study, WM refers to both storage and processing of information, measured by reading or listening span tasks (Fortkamp, 1999; van den Noort et al., 2006). Phonological short-term memory refers to storage alone, measured by nonword repetition or nonword recognition (van den Noort et al., 2006).

Working memory is one of the most extensively investigated factors relating to individual differences in cognition. Baddeley's model of WM (Baddeley, 1998, 2003; Baddeley & Hitch,

1974), comprised of the central executive, the phonological loop, and the visuo-spatial sketchpad, is one of the most influential. Baddeley (2003; Baddeley, Gathercole, & Papagno, 1998) has proposed that the function of the phonological loop is to support language learning, including vocabulary development, fluency, and some measures of comprehension (e.g., Ellis, 2006; French, 2006; Service, 1992). Central executive function has also been shown to relate to reading comprehension and global verbal abilities (e.g., Daneman & Carpenter, 1980; Turner & Engle, 1989). It may also account for individual differences in the efficiency of language processing, both in first language (L1) and L2 (Harrington & Sawyer, 1992; Williams, 2012).

### *PSTM and L2 Learning*

A number of studies have shown that PSTM is important in L2 vocabulary learning. Baddeley, Papagno, and colleagues (Baddeley, Papagno, & Vallar, 1988; Papagno, Valentine, & Baddeley, 1991) studied a neuropsychological patient, PV, who had a selectively damaged phonological loop. She learned pairs of native language words without difficulty, but was greatly impaired when attempting to learn foreign language words. The importance of the phonological loop was further supported by Papagno et al. (1991) who showed that normal participants were unable to learn foreign language words under articulatory suppression, an interference treatment that selectively affects PSTM, especially when the novel words were very dissimilar from those of their native language.

In addition to vocabulary, PSTM has also been implicated in L2 grammar and fluency development. Service (1992) and Service and Kohonen (1995) conducted a longitudinal study of foreign language learning in Finnish-speaking primary school children. This study found that English (L2) nonword repetition abilities at the beginning of primary education were a good predictor of success in English learning during the first 2-3 years of formal education (Service,

1992). Ellis and Sinclair (1996) tested adults' ability to learn Welsh as a foreign language and found that participants who repeated the language aloud scored significantly higher on vocabulary, use of phrasal constructions, and the ability to use the Welsh soft mutation whereby the initial phonemes of nouns change as a function of grammatical context (King, 2003). They concluded that the more often foreign language structures are rehearsed in PSTM, the easier it is to learn them and to generalize rules from them. Kormos and Sáfár (2008) found similar results: that nonword span correlated with writing, use of English (both vocabulary and grammar constructions) and overall L2 proficiency.

PSTM is important for the development of overall L2 fluency in addition to vocabulary and grammatical structures. French and O'Brien (2008) found that Time 1 nonword repetition predicted L2 grammar scores at Time 2 ( $r$ s between .79-.82), after a 5-month intensive language program for children. Phonological memory explained almost 30% of the variance in grammar scores at Time 2, even after controlling for vocabulary knowledge. O'Brien and colleagues also found important influences of PSTM on measures of adult L2 learning, including vocabulary, correct use of grammatical structures, and fluency (O'Brien, Segalowitz, Collentine, & Freed, 2006; O'Brien, Segalowitz, Freed, & Collentine, 2007). O'Brien et al. (2006) found that phonological memory correlated with vocabulary scores, narrative abilities, and use of free grammatical morphemes and subordinate clauses, both at the beginning and at the end of a semester of Spanish learning ( $r$ s between .30-.41). O'Brien et al. (2007) extended these findings to measures of overall fluency in adult L2 learners.

Speidel (1993) studied siblings who were native speakers of German and L2 speakers of English. The sibling who had trouble with gender forms and case endings also had trouble with PSTM, and Speidel concluded that PSTM is important in the creation of stable representations of

grammatical structures. Difficulties in creating these representations may lead to problems building a storehouse of token phrases from which to generalize grammatical rules. Speidel suggested that PSTM would be especially important for learning material not easily learned through meaningful imagery, such as abstract function words and morphemes (DeKeyser, 2005).

Williams and Lovatt (2003) investigated the relationship between PSTM and the ability to generalize grammatical gender. Using a semiartificial foreign language (the structure of the language was that of Italian, but with the words changed to Japanese nonwords and the letters of the determiners also changed), they found that phonological memory predicted grammatical generalization abilities at  $r = .60$  and that this correlation was significant across multiple cycles of generalization tests. Similar to Ellis (2006), Williams and Lovatt argued that grammar rules are generalizations of patterns across sequences of words. They reasoned that if PSTM is related to learning words, it should also be related to learning the patterns among them. They suggested that because there appears to be a consistent relationship between PSTM and vocabulary, and between vocabulary and grammar, it is reasonable to assume a connection between PSTM and grammar as well.

### *WM and L2 Learning*

Working memory-impaired children have difficulties parsing and analyzing linguistic structures in their L1 (Marton & Schwartz, 2003; Robinson, Mervis, & Robinson, 2003). Research has also implicated WM in L2 comprehension, reading, and fluency. Harrington and Sawyer (1992) used a L2 reading span test based on Daneman and Carpenter (1980), simplified for use with nonnative English speakers. They found that this reading span task correlated with both L2 reading and grammar scores. Fortkamp (1999) replicated these results and also developed a new WM task: speaking span. In this task, participants were briefly presented with a

list of words and then had to generate a set of sentences, with one sentence each incorporating one of the words. Scores on this task also correlated with scores of L2 fluency ( $r$ s between .61 and .64). Leiser (2007) found that participants with higher WM capacities were better able to comprehend passages in a foreign language (although only when the participants were familiar with the passage topic). He also found that having a high WM capacity compensated for being unfamiliar with a passage topic on a test of grammatical form recognition.

In addition to their PSTM results, Kormos and Sáfár (2008) found that WM (measured using backward digit span) correlated with five out of their six measures of L2 ability, including reading, speaking, and listening. Their findings were important and unique because they demonstrated a correlation not only between WM and typical verbal abilities such as reading and listening comprehension but also between WM and vocabulary knowledge as assessed by the Cambridge First Certificate Exam. They argued that this demonstrated a direct connection between WM and both vocabulary and grammar learning that is based on the role of the central executive component of WM in regulating attention (Baddeley, 2003). Working memory has similarly been proposed as a basic mechanism for learning new rules in a L2 through its involvement in the noticing and encoding of new information (Mackey, Philp, Egi, Fujii, & Tatsumi, 2002). Working memory is also important for L2 grammar learning across time. Robinson (2002) used a reading span task to measure WM and found it was related to the incidental learning of Samoan grammatical structures. Participants with higher WM scores performed better on sentence production and receptive grammaticality judgment tasks. This was true not only on immediate posttests ( $r = .42$ ) but also on 1-week ( $r = .33-.48$ ) and 6-month ( $r = .44$ ) delayed posttests.



Sunderman and Kroll (2009) used a reading span task to investigate the WM capacities of college students who studied abroad and those who did not. Working memory scores correlated significantly with performance on a translation comprehension task. However, they also found that beneath a minimum WM capacity threshold, students who studied abroad were not able to benefit from that experience. They concluded that individuals with higher WM may be able to attend to more linguistic factors at once and thereby increase their ability to parse grammatical structures. As a whole, this research demonstrates the important role of WM in various domains of L2 learning.

#### *Summary of Previous Findings*

A large amount of research has indicated that both PSTM and WM are important for various aspects of language learning ability. Most researchers attribute the connection between PSTM and language learning to the importance of the phonological loop for forming stable, long-term mental representations of novel phonological material. These representations are especially important for knowledge of phonological items, such as individual words and chunks. Connections have also been found between WM and language abilities. These relationships are usually attributed to an individual's ability to parse, analyze, and effectively manipulate new linguistic items and structures. The attentional aspect of WM has additionally been implicated in allowing learners to attend to multiple aspects of linguistic structure at once.

However, there are still many unanswered questions. Much of the work that investigates PSTM and WM in grammar learning has involved explicit learning conditions, with relatively little research done on implicit or naturalistic learning conditions. Additionally, some authors have suggested that the relationship between PSTM and grammar abilities is mediated by vocabulary knowledge (e.g., Service & Kohonen, 1995). A large amount of research has been

correlational or observational, rather than experimental; further work is necessary to more clearly identify the separate roles of PSTM and WM in vocabulary and grammar learning. The current study intended to address these issues, with a particular focus on uninstructed grammar learning where patterns are induced from whole sentences.

### The Current Study

Three memory tests were used: listening span to measure WM (Harrington & Sawyer, 1992), and nonword repetition (Gathercole, Pickering, Hall, & Peaker, 2001) and nonword recognition (O'Brien et al., 2006, 2007) to measure PSTM. An artificial language utilizing nonword stimuli was employed for the language learning tasks (similar to Williams & Lovatt, 2003). Participants learned the vocabulary and were then exposed to word order and plural markings in a sentence context, without explicit instructions or explanations. Their knowledge and use of plural markings and word order was measured in a generalization test at the end of the study; these scores reflected not only their knowledge of the structures but also their ability to generalize them to new words and phrases. Three specific hypotheses guided the current study:

1. There will be a positive correlation between PSTM and vocabulary scores. This is based on findings, such as Baddeley et al. (1988), of a strong connection between phonological memory and L2 vocabulary learning.
2. There will be a positive correlation between both PSTM and WM and grammar scores. This is based on the fact that grammar learning is a complex process, and relies on both memorization of individual items and processing of the relationships between them.

3. There will be a positive correlation between vocabulary and grammar scores. This is based on prior research that demonstrates a relationship between vocabulary knowledge and grammar abilities in L1 (Bates & Goodman, 1997) and L2 (Service & Kohonen, 1995).

## Method

### *Participants*

Fifty monolingual native English speakers were recruited from a large American university in the Midwest. Three participants were excluded because they grew up as bilinguals; there were 7 additional participants who did not return for the second session. This left a total of 40 participants (36 females, 4 males) who completed the entire study and whose data were included. Ages ranged from 18 to 45, with a mean of 21.5 years. Participants volunteered and were paid \$10.00 per session, for a total of \$20.00.

### *Individual Differences Measures*

*Nonword Repetition.* Nonword repetition was used as a test of PSTM. Participants heard a list of one-syllable nonwords and were asked to repeat them as accurately as possible. There were 4 lists at each of 4 lengths: 3, 4, 5, and 6 words. All participants heard the lists in the same order, beginning with the shortest lists and increasing in length. The nonwords were taken from Gathercole et al. (2001). Examples of these stimuli can be found in Appendix A. Participants' responses for all items throughout the study were recorded using a microphone and the sound-editing software Audacity (Audacity Team, 2008). Scoring was done offline on a phoneme-by-phoneme basis and the maximum number of phonemes recalled on any one repetition set was

calculated for each participant. The maximum possible score for this task was 22 correct phonemes.

*Nonword Recognition.* Nonword recognition was used as an additional measure of PSTM. It is highly correlated with performance on similar repetition tasks, but is less affected by unfamiliar phonotactics and pronunciation difficulties (Gathercole et al., 2001). Participants listened to two presentations of a list of nonwords and decided whether they were the same or different. Participants received 1 point for each correct same or different judgment. Eight lists were used at each of three lengths: 5, 6, and 7 items. There were also four practice trials with 4 items each. Of the eight lists at each length, four were identical and four were different. For the identical presentations, the same list was presented twice with a 1200 ms pause in between. For the presentations that were different, the first presentation of the list was followed by a 1200 ms pause and a second presentation of the list with two adjacent items transposed. The location of the transposed syllables was randomized, with the exception that the first and last syllables were never transposed. This was to reduce the salience of the transposition and encourage participants to process the entire string (O'Brien et al., 2006). The stimuli were taken from Gathercole et al. (2001) and also used by O'Brien et al. (2006). Examples of these stimuli can be found in Appendix A. The maximum possible score for this task was 24 correct recognitions.

Participants began with the practice set of four lists, on which they received feedback (pilot-testing revealed this was necessary for learning the task). Test trials began with the 5-item lists and then moved on to the 6- and 7-item lists, with no feedback. Within each list length, the order of the same and different lists was randomized for each participant. All nonwords used as vocabulary for the artificial language were used as stimuli in this task. This was done to

familiarize participants with the phonotactics of the words in the language they would be learning (Brooks, Braine, Catalano, Brody, & Sudhalter, 1993; Williams & Lovatt, 2003).

*Listening Span.* Listening span, adapted from similar reading span tasks (e.g., Daneman & Carpenter, 1980; Harrington & Sawyer, 1992), was used to measure WM capacity. Participants heard sentences and had to decide whether or not they made sense as a sentence in English. This is referred to as a grammaticality judgment, even though the judgments were not made on strict grammaticality per se, but rather on whether the sentences made sense. The ungrammatical sentences were clearly so because their word orders made them completely nonsensical. This difference was explained to participants. At the end of each set of sentences, participants were asked to recall the final word of each sentence. The number of words correctly recalled in the correct relative order was used as the WM score; the maximum possible score was 31. The grammaticality judgments were not scored, but served as a manipulation check to ensure that the participants processed each sentence as a whole (Turner & Engle, 1989). All participants were correct on at least 85% of the grammaticality judgments and all trials were included in calculating the WM scores. Forty sentences were used and arranged in sets, with the number of sentences in each set increasing from 2 to 6. Two sets of each length were presented. The two sets of two sentences were practice sets; the longer sets were used as test sets. The sentences used as stimuli were from Harrington and Sawyer (1992); examples can be found in Appendix A. Half the sentences were grammatical and half the sentences were ungrammatical (the word order was mixed up so that it no longer made sense as a sentence in English).

#### *The Artificial Language*

The artificial language consisted of 21 nouns, 10 verbs, 4 adjectives, and 3 prepositions. Each noun and verb had three forms: singular, dual, and plural, marked by a prefix. Nouns had

no prefix for the singular form, the prefix *zi-* for the dual form, and the prefix *na-* for the plural form, with one sound change in irregular forms. Nouns that began with nasal consonants (/m/ or /n/) took the prefix *za-* for the dual form and the prefix *no-* for the plural form. Verbs had no prefix for the singular form, the prefix *ta-* for the dual form, and the prefix *moo-* for the plural form, also with one sound change in irregular forms. Verbs that began with nasal consonants took the prefix *too-* for the dual form and the prefix *mi-* for the plural form. Adjectives took the same prefixes as the nouns they modified. A complete set of words in all forms can be found in Appendix B. The word order of the artificial language was the same as English, except that adjectives followed the nouns they modified, rather than coming before them.

Sentence (1) demonstrates a singular sentence. Sentences (2) and (3) demonstrate plural forms, where DF indicates the dual prefix. Sentence (2) demonstrates regular prefixes and (3) demonstrates their irregular forms.

(1) *Lork cham mord kib dook*

Cow big is on table

“The big cow is on the table.”

(2) *Na- targ moo-dern zi-jick zi-leck*

plural-fish plural-throw DF-book DF-red

“Fishes throw two red books.”

(3) *No-nog moo-pag za-nart*

plural-man plural-catch DF-ball

“Men catch two balls.”

Although it was an artificial language, the language used here was designed on the basis of both other experimental language systems (such as in Daneman & Case, 1981) and attested

natural languages. Although the usage of these structural features are to a certain extent simplifications of their forms and functions in natural languages, as examples Arabic uses a singular-dual-plural distinction such as the one used here (Haywood & Nahmad, 1965), and Swahili uses prefixes to mark various classes of words (Polomé, 1967).

### *Procedure*

The study was completed over the course of two 1-hr sessions in a computer lab. All tasks, except for the participant questionnaires—which were completed on paper—were administered using E-Prime software (Schneider, Eschman & Zuccolotto, 2002). During the first 1-hr session, participants completed all three memory tasks and then began learning the artificial language. Participants progressed through the rest of the study at their own pace. The vocabulary for the language was presented aurally through headphones with the corresponding meaning illustrated on the screen. Participants were asked to repeat each word aloud to reinforce its learning (Ellis & Sinclair, 1996). They also heard the translation and were asked to repeat that as well. The singular form of each noun and verb was presented twice. The dual and plural forms of each word were never presented in this phase. Participants never saw the foreign language in written form.

After presentation of the vocabulary, participants listened to 15 sentences in the foreign language that contained this same vocabulary. They also saw an illustration of the sentence and its English translation on the screen and were asked to repeat the sentence (repetition paradigm). An example illustration for the sentence *The man throws the egg* is provided in Figure 1. After this initial presentation, participants listened to the same 15 sentences a second time. This time, they saw the illustration, heard the English sentence, and were asked to translate into the foreign language (translation paradigm). After each sentence, they pressed a key and heard the correct

translation as feedback. These sentences were presented in random order and some examples can be found listed in Appendix C under Set 1. Vocabulary scores from this session are included in the analyses as Initial Vocabulary scores. This concluded the first session.

---

INSERT FIGURE 1 ABOUT HERE

---

During the second session, 2 to 7 days after the first, participants began with a vocabulary review. They heard a word in the foreign language and saw four illustrations on the screen. They had to choose which illustration matched the word they heard. They were given feedback on whether their answer was correct or incorrect, and the word and its illustration were repeated to reinforce the meaning. This review continued until participants reached 85% correct on a single run-through of all 38 words. After this, participants reviewed the 15 sentences from session 1 (Set 1). Half the sentences were presented using the repetition paradigm and the other half were presented using the translation paradigm. The sentence order within each half was randomized for each participant.

After the review, participants were presented with 30 entirely new sentences (examples are listed in Appendix C under Set 2). For the first half, participants saw an illustration of the sentence and its English translation. They heard the sentence in the foreign language and were asked to repeat it (repetition paradigm). For the second half, they saw an illustration, heard a different sentence in English, and were asked to translate it into the foreign language (translation paradigm). After their production they heard the correct translation as feedback. Within each half of the stimuli, sentences were presented in a random order for each participant. Twenty-three of these new sentences included at least one unknown plural form of a known vocabulary



word. These plural forms had not been previously encountered and were not explicitly taught. The meaning of these new forms could only be deduced from the illustration and the feedback that accompanied each sentence. These provided the data from which participants generalized the grammar rules.

The final phase of the study was the test productions (TP) phase, consisting of 50 sentences. Thirteen sentences were repeated from Set 1, 13 were repeated from Set 2, and 24 were completely novel (examples are listed in Appendix C under Generalization Set; GS). The novel sentences of the GS each used at least one new plural form that participants had not yet encountered. For half the sentences, participants heard an English sentence, saw an illustration, and were asked to produce the foreign language translation (production scores). For the other half of the sentences, participants heard an utterance in the foreign language and had to translate into English (comprehension scores). On these trials, participants did not see an illustration because this would have allowed them to translate the sentence solely on the basis of the illustration, without necessarily having understood the foreign language. Participants received separate scores for production and comprehension of GS vocabulary and grammar as well as overall TP vocabulary and grammar.

The purpose of the novel sentences was to measure how well participants had abstracted the plural morphology and word order rules from the sentences presented in Set 2 and could then generalize them to new vocabulary and sentences. This served as a measure of their knowledge of the grammar system. At the end of the second session, participants completed a questionnaire that assessed their explicit knowledge of the language. Although the language was presented with an implicit learning orientation, it is likely that participants attempted to analyze the system explicitly and did not simply rely on implicit learning. Therefore this questionnaire was used to

evaluate how much explicit knowledge was gained through the implicit or naturalistic learning environment. The participants were asked to describe the grammar of the language, to give the rule for forming plurals, and to identify whether there were any apparent changes to these rules in any contexts. They were also asked for any additional comments they might like to share regarding how they learned the language.

### *Scoring*

Utterances were scored for both vocabulary and grammatical accuracy. For vocabulary production scores, participants received 1 point for each word that was produced correctly, .66 points for each word that had only one incorrect phoneme, or 0 points for each word that had two or more incorrect phonemes. For example, if the target word was *charb*, participants would earn 1 point for producing *charb*, .66 points for producing *chard*, and 0 points for producing *chob*. For grammar production scores, participants received 1 point for each correct prefix, .25 points for each prefix that was correct in word type (noun, verb, or adjective) but incorrect in plurality (dual vs. multiple), or 0 points for a plural marking that was missing or from the wrong word type. For example, if the target word was *zi-charb*, participants would earn 1 point for producing *zi-charb*, .25 points for producing *na-charb*, and 0 points for producing *charb* or *ta-charb*. Participants also received 1 point if they were able to produce the correct noun-adjective word order. This partial scoring scale was used to give participants credit for partially learning the vocabulary and morphology system, and similar criteria have been used by other researchers (e.g., Williams & Lovatt, 2003).

For vocabulary comprehension scores, participants received one point for each word they correctly translated into English. For grammar comprehension scores, participants received one point for each plural distinction they correctly expressed in English, .25 points if they expressed

a multiple-noun meaning but chose the wrong one, or 0 points if they expressed a singular meaning when it should have been a multiple meaning (or vice versa). For example, if a participant produced *Many babies eat an apple* when the target production was *Two babies eat two apples*, he or she would receive 3 points (full credit) for vocabulary and .25 points for grammar (.25 for *two* instead of *many* and 0 for *an* instead of *two*). Scores are labeled according to which stimuli set they belong to: TP (Test Productions) scores are the composite scores from the final test phase, and combine comprehension and production of both old and new stimuli. Generalization Set (GS) scores are those earned only on the novel stimuli from the final test phase. For these scores, both the composite scores (just labeled GS) and scores separated out for production and comprehension are reported. Initial Vocabulary scores are taken from performance on the Set 1 sentences during the first session. Vocabulary scores consist of accuracy in the vocabulary items, and grammar scores consist of accuracy in the morphology as well as the word order of the language.

Participants' foreign language learning experience was assessed by calculating the total number of years spent studying foreign languages (mean = 6.95, SD = 3.87, range 1-16.5 years). Responses to the questionnaire assessing participants' explicit knowledge of the new language were converted to numeric data by a point scale. Participants received one point for each unique aspect of the grammar they listed in their description of the overall grammatical structure (labeled Describe Grammar) and one point for each aspect of the rule for forming plurals or specific example that they listed (labeled Rule for Plurals). For example, a participant would receive one point each for saying "adjectives come after the noun" and "plurality is indicated by prefixes" in the *Describe Grammar* section. A participant would also receive one point each for listing the plural prefixes *no-* and *na-* and labeling them as such.

## Results

Descriptive statistics for all major variables are listed in Table 1. Figure 2 shows the mean percent correct for vocabulary and grammar scores at various time points during the experiment. Although these average percentages are not high, the extremes of the range of results (listed in Table 1) indicate that some participants were indeed able to learn the language quite accurately. Because of the nature of the production task, it is not possible to compare these results to chance. Nevertheless, from fewer than two hours of exposure, the amount of learning demonstrated by the learners is quite substantial.

---

INSERT TABLE 1 ABOUT HERE

---



---

INSERT FIGURE 2 ABOUT HERE

---

### *Correlational Analyses*

*Memory Measures and Vocabulary Scores.* Correlations between the memory measures and the vocabulary scores are presented in Table 2. The PSTM and WM measures do not correlate with each other; this indicates that they measured distinct cognitive abilities.

---

INSERT TABLE 2 ABOUT HERE

---

Nonword repetition correlated with GS vocabulary scores, both in production ( $r = .33$ ) and comprehension ( $r = .42$ ). Accuracy on the nonword recognition task also correlated with comprehension of vocabulary in the GS ( $r = .45$ ). Overall, these results indicate a significant

relationship between phonological memory and vocabulary knowledge during the final phase of the study and are consistent with the previous literature (Baddeley, Papagno, & Vallar, 1988; Masoura & Gathercole, 1999).

Working memory also correlated with production of vocabulary in the GS,  $r = .34$ . However, it did not correlate with vocabulary comprehension. The finding of a correlation between WM and vocabulary performance is somewhat surprising because it is not commonly reported, although it is not completely without precedent (e.g., Kormos & Sáfár, 2008).

*Memory Measures and Grammar Scores.* The main purpose of this study was to explore the relationship between PSTM and WM and grammar learning in an artificial foreign language. To do this, correlations were computed between PSTM and WM and various measures of grammatical ability (including both morphology and word order) from the end of the study. These correlations can be found in Table 3. Nonword repetition correlated with all three grammar scores, from both TP and the GS ( $r$ s from .34 to .43). Nonword recognition accuracy did not correlate with any measures of grammar ability. Working memory correlated with grammar abilities slightly more strongly than did PSTM, with correlations between WM and all three composite grammar scores ranging from  $r = .35$  to  $r = .46$ . Working memory, but not PSTM, also correlated with participants' scores for describing the rule to form plurals ( $r = .33$ ). These results indicate that both PSTM and WM have strong relationships with participants' ability to generalize and apply grammar rules in both production and comprehension.

---

INSERT TABLE 3 ABOUT HERE

---

*The Relationship between Vocabulary and Grammar.* A strong relationship between vocabulary knowledge and grammar abilities has been found in the previous literature (e.g.,

Bates & Goodman, 1997; Service & Kohonen, 1995). To determine whether this relationship appeared in the present study, correlations between measures of vocabulary and grammatical ability were calculated and can be found in Table 4. Overall, the relationship between vocabulary and grammar abilities was very strong, with  $r$  values ranging from .44 to .76. A series of hierarchical multiple regression analyses were therefore conducted to determine which individual factors contributed to variance in language ability in this study.

---

INSERT TABLE 4 ABOUT HERE

---

### *Regression Analyses*

*Memory and Vocabulary.* Two hierarchical multiple regression analyses were performed with the Initial Vocabulary scores (performance on vocabulary during the training of Set 1 in the first session) and the GS composite vocabulary scores (production and comprehension combined, from the novel sentences during the final test phase) as the dependent variables in order to identify whether PSTM and WM play independent roles in vocabulary learning. The memory measures were used as predictor variables. Details of these analyses can be found in Table 5. For the Initial Vocabulary score, age and years of language study were forced into the model first to control for their influence, but did not account for any significant variance in the vocabulary scores. When nonword repetition was entered into the model, it explained a significant 16% ( $\beta = .43$ ) of the variance in vocabulary scores. Working memory contributed a further 17% ( $\beta = .43$ ) of the variance. For the GS composite vocabulary score, age and years of language study were again entered into the model first but accounted for no significant variance. When nonword repetition was added to the model, it predicted a significant 14% ( $\beta = .39$ ) of the variance in

vocabulary scores. Working memory accounted for an additional 10% ( $\beta = .35$ ) of the variance. These results demonstrate that PSTM and WM have significant independent influences on vocabulary abilities at multiple beginning stages of language learning.

---

INSERT TABLE 5 ABOUT HERE

---

*Memory and Grammar.* Two hierarchical multiple regression analyses were conducted to determine the relative influences of PSTM and WM on foreign language grammar scores. The dependent variables were GS grammar production and comprehension, and the predictor variables in each analysis were the three memory measures. Detailed results of these analyses are presented in Table 6. For grammar production, age and years of language study were forced into the model first to control for their influence. Although the beta coefficient for years of language study was significant, when combined together in one step, age and years of study failed to account for a significant amount of variance in grammar production scores. Working memory was entered in the next step and explained a significant 14% ( $\beta = .40$ ) of the variance. Nonword repetition contributed an additional significant 10% ( $\beta = .34$ ) of the variance in the third step. For grammar comprehension, age and years of language study were again entered but accounted for no significant variance. Nonword repetition was entered in the next step and explained a significant 17% ( $\beta = .43$ ) of the variance, and WM contributed an additional significant 11% ( $\beta = .36$ ) of the variance. These analyses supported the finding that PSTM and WM each separately contribute a significant amount of variance to the ability to learn, generalize, and produce correct grammatical structures in a foreign language.

---

---

 INSERT TABLE 6 ABOUT HERE
 

---

*Are Effects of PSTM and WM upon Grammar Independent or Mediated by Vocabulary?*

Some researchers have suggested that the relationship between memory measures and grammatical competency is mediated by vocabulary (e.g., Service & Kohonen, 1995). There are two different ways to operationalize this question. The first involves testing whether there are unique contributions of the memory measures to final attainment grammar over and above the variance explained by final attainment vocabulary. The second concerns whether the contributions of the memory measures to final attainment grammar are entirely mediated by initial vocabulary learning, or whether there are separable memory involvements in inducing grammar above and beyond knowledge of the component lexis. This second analysis relates to the proposal of Service and Kohonen (1995) that PSTM does support grammar learning, but that it does so via effects on vocabulary learning rather than directly on grammar learning per se.

To test the first hypothesis, further regression analyses were conducted with both GS grammar production and comprehension as dependent variables. The final state vocabulary score from the GS was forced in as the first predictor variable to control for its influences. This vocabulary score accounted for a significant amount of variance for both grammar variables: production  $Adjusted R^2 = .55$ ,  $F(1, 38) = 48.43$ ,  $p < .001$ ,  $\beta = .75$ , 95% CI [.54, .96]; and comprehension  $Adjusted R^2 = .29$ ,  $F(1, 38) = 17.16$ ,  $p < .001$ ,  $\beta = .56$ , 95% CI [.29, .82]. None of the PSTM or WM measures explained any additional variance once vocabulary had been entered into the models. This is in contrast to the previous analysis, which did not consider vocabulary knowledge, and in which WM accounted for a significant 14% ( $\beta = .40$ ) of the variance in



production and PSTM a significant 10% ( $\beta = .34$ ). Additional analyses demonstrated that age and foreign language study also did not contribute any unique variance in grammar scores.

A second regression analysis was conducted on the TP grammar scores (performance on both repeated and novel stimuli). Again, age and years of language study were entered first and did not account for any significant variance. Test Productions vocabulary was entered next to control for its influences and explained a further 48% of the variance in grammar ( $F(3, 36) = 41.49, p < .001, \beta = .72, 95\% \text{ CI } [.50, .94]$ ). In this case, however, the addition of WM explained an additional 4% of the variance in grammar scores, beyond that accounted for by vocabulary knowledge (total *Adjusted R*<sup>2</sup> = .55,  $F[4, 35] = 24.77, p < .001, \beta = .24, 95\% \text{ CI } [.02, .46]$ ). In contrast, when vocabulary knowledge is not considered, WM accounts for 19% ( $\beta = .46$ ) of the variance in overall TP grammar scores, and PSTM accounts for an additional significant 12% ( $\beta = .37$ ). These analyses indicate that although the relationship between PSTM and WM and grammar abilities is shared with vocabulary knowledge, WM does still explain a significant additional amount of variance in grammatical abilities beyond that explained by vocabulary.

---

INSERT FIGURE 3 ABOUT HERE

---

The second operationalization relates to the Service and Kohonen (1995) proposal regarding PSTM and grammar learning. Service and Kohonen state that PSTM is involved in the initial learning of vocabulary, and that vocabulary is also involved in learning grammar. Therefore, PSTM is involved in learning grammar, but only indirectly so. To test this hypothesis, causal path analyses were performed using AMOS (Arbuckle, 2006) to determine whether there are independent effects of PSTM and WM upon ultimate (GS) vocabulary and grammar that are not mediated by explicitly learned vocabulary at the end of the training phase. The causal path

model shown in the left panel of [Figure 3](#) investigates the direct effects of PSTM and WM on vocabulary scores at the end of the training session, the direct effects of trained vocabulary upon GS vocabulary and grammar, and whether there remain any additional independent direct effects of PSTM or WM on GS vocabulary and grammar. Phonological short-term memory and WM are the exogenous variables; the others are endogenous and as latent variables have their associated errors, labeled as  $e$  in the figure. The proportion of nonerror variance explained is shown for each endogenous variable (e.g., the model explains 53% of the variance in GS vocabulary). The pathweights for all of these effects are shown in the left panel of [Figure 3](#). Some of these paths were significant, some not. A second model was run which specified just these significant paths and recalculated the effects as shown in the right panel of [Figure 3](#). This demonstrates (a) substantial independent effects of PSTM (.41) and WM (.42) upon trained vocabulary, (b) a large autocorrelation between trained vocabulary and later GS vocabulary (.72), (c) an effect of trained vocabulary upon GS grammar (.33)—which in turn mediates indirect effects of PSTM (.14) and WM (.14) upon GS grammar (i.e.,  $0.41 \times 0.33$  and  $0.42 \times 0.33$ , respectively)—and (d) direct effects of PSTM (.25) and WM (.30) upon GS grammar. These analyses suggest that there are indeed significant independent effects of PSTM and WM upon L2 vocabulary learning and upon L2 grammar learning, some of which are mediated by vocabulary and some of which are direct effects.

#### *What Predicts the Emergence of Explicit Meta-linguistic Knowledge?*

Even though the primary purpose of this study was to explore the connections of both PSTM and WM with learning from naturalistic exposure, it is also possible that PSTM and WM are involved in the emergence of explicit grammar knowledge in this type of learning condition. To explore this possibility, a number of additional analyses were performed. First, correlations

were calculated between the memory measures, initial and final vocabulary scores, final grammar scores (both production and comprehension), and participants' scores on two measures of explicit knowledge taken at the end of the study: Describe Grammar and Rule for Plurals (see the Scoring section for descriptions of these variables). Describe Grammar correlated with the vocabulary scores with  $r$ s ranging .33-.42, but did not correlate with any grammar scores. Rule for Plurals correlated with vocabulary scores with  $r$ s ranging .41-.60, and also correlated with grammar scores with  $r$ s ranging .55-.61. Details of these correlations can be found in Table 7.

---

INSERT TABLE 7 ABOUT HERE

---

Given the fact that a number of significant correlations were found, regression analyses were used to further investigate explicit knowledge scores. These analyses focused on the Rule for Plurals measure, as it had the strongest relationship with the other variables. The first analyses explored explicit knowledge as a potential predictor of grammar learning, with GS grammar production and comprehension as the dependent variables. Age and years of language study were entered first to control for their influence, but did not account for any significant variance. Rule for Plurals was entered in the next step and accounted for an additional 23% ( $\beta = .49$ ) of variance in grammar production scores (*Adjusted R*<sup>2</sup> = .30,  $F(3, 36) = 6.43$ ,  $p < .001$ ). The memory measures did not account for any additional variance in the grammar production scores. For comprehension scores, Rule for Plurals accounted for a significant 24% ( $\beta = .52$ ) of variance (*Adjusted R*<sup>2</sup> = .26,  $F(3, 36) = 5.64$ ,  $p < .01$ ). This time, however, nonword repetition (a PSTM measure) accounted for an additional 11% ( $\beta = .35$ ) of variance in final grammar comprehension scores (*Adjusted R*<sup>2</sup> = .37,  $F(4, 35) = 6.76$ ,  $p < .001$ ).

Given the clear relationship between explicit grammar knowledge and final vocabulary and grammar scores, the additional analysis focused on accounting for variance in the explicit knowledge scores themselves. Rule for Plurals was used as the dependent variable and the predictor variables were the memory measures and GS vocabulary and grammar scores. Age and total years of study were entered first but did not account for any significant variance. In the next step, GS vocabulary was entered and accounted for a significant additional 31% ( $\beta = .57$ ) of the variance (*Adjusted R*<sup>2</sup> = .32,  $F(3, 36) = 7.06$ ,  $p < .001$ ). The memory measures and final grammar scores did not contribute any additional variance.

Even though the results of these analyses are interesting, they must be interpreted with extreme caution for the following reasons: (a) although they were analyzed directionally in these two regressions, the patterns are no more than correlations because growth in all of these variables was taking place simultaneously over the second session of the experiment; (b) all of the explicit knowledge measures were taken at the end of the study and were administered using open-ended questions such as “Tell me a little bit about the grammar of the language you just learned” on which overall scores were extremely low (an average of only 1.2 points, and a mode of 1 point for this question). Because of the limitations of these data and how they were collected and measured, these results are no more than exploratory. However, this is clearly an area where future research may prove fruitful.

## Discussion

The goal of the present study was to explore the relationship between PSTM and WM and the ability to learn grammatical patterns in a foreign language. Participants completed three memory tests: nonword repetition, nonword recognition, and listening span. They learned

singular vocabulary words and sentences in a foreign language and were then exposed to plural forms in a sentence context with no instruction on these novel forms. They were then tested on their production and comprehension of fifty sentences which included novel plural expressions. Measures of their language abilities at the end of the study were used as the dependent variables.

### *Memory Measures and Vocabulary*

As predicted, positive correlations between PSTM and final vocabulary were found. The magnitude of these correlations, generally .33-.45, is moderately strong and is consistent with previous findings (e.g., Masoura & Gathercole, 1999). Both nonword repetition and nonword recognition correlated with vocabulary production and comprehension. Working memory also correlated with vocabulary production, though not comprehension. This result was less expected because WM is usually associated with variables such as reading comprehension and fluency rather than vocabulary (Daneman & Carpenter, 1980; Harrington & Sawyer, 1992; Sunderman & Kroll, 2009; Turner & Engle, 1989; but see Kormos & Sáfár, 2008), albeit the strength of the correlations between WM and vocabulary were weaker than between PSTM and vocabulary.

The regression analyses confirmed that PSTM and WM make independent significant contributions to vocabulary learning. After the influences of age and foreign language study were accounted for, PSTM accounted for 14% ( $\beta = .39$ ) of the variance in final vocabulary scores and WM accounted for an additional 10% ( $\beta = .35$ ). It seems, therefore, that PSTM and WM are two related but separable memory constructs that make independent contributions to vocabulary. Although the actual mechanisms underlying the influence of PSTM and WM are still not fully understood, previous research on these constructs allows for a hypothesis to be put forth. Phonological short-term memory likely supports the consolidation of stable phonological representations in long-term memory (Ellis, 1996, Jones, Gobet, & Pine, 2007, 2008). As an

attentional control system, WM may support the maintenance of relevant information and the regulation of processing during complex operations, with any spare capacity allowing for the noticing of novel task-relevant features and their integration into the system (Mackey et al., 2002).

### *Memory Measures and Grammar*

Previous research has reported correlations between PSTM and L2 grammatical ability and fluency with  $r$  values ranging from .30-.80 (Adams & Gathercole, 1996; 2000; Daneman & Case, 1981; French, 2006; Service, 1992) and between WM and L2 grammatical abilities and fluency with  $r$  values ranging from .30-.65 (Fortkamp, 1999; Harrington & Sawyer, 1992; Kormos & Sáfár, 2008). The hypothesis of a relationship between the memory measures and grammar learning was also supported in the present study. A number of significant correlations were found between measures of PSTM and final grammar scores, most with strengths of  $r = .30-.45$ . This replicates previous findings of a relationship between PSTM and grammar abilities (e.g., Kormos & Sáfár, 2008; O'Brien et al., 2006, 2007; Williams & Lovatt, 2003). The strengths of the relationships between PSTM and grammar and between PSTM and vocabulary were remarkably similar, underscoring the importance of PSTM for both language domains.

Working memory also correlated with final grammar scores, and the correlations between WM and grammar production were consistently stronger than between PSTM and grammar. This finding was not unexpected. Working memory, by definition, includes both storage and processing of information (Baddeley, 1998, 2003). The storage component of WM explains its relationship with vocabulary learning—that is, remembering individual items. Grammar learning, however, depends on much more than just memorizing items. It is also the process of abstracting patterns from across language sequences presented as input (Ellis, 1996; Speidel,

1993). Vocabulary learning involves the sound patterns of words and their arbitrary mapping to meaning; grammatical patterns involve abstracting the relations between vocabulary items and identifying their functional significance. Grammatical patterns are more global and apply to the utterance as a whole, not just the individual word. These more complicated patterns therefore demand more processing capacities, the holding of a greater amount of information over time, and the identification, selection, and correlation of relevant features both in the input and in long-term memory. These are the aspects provided by WM over PSTM.

Using grammatical knowledge to produce novel utterances also requires greater processing. Once learners have understood the input and parsed the utterance into individual units, they must be able to process those units and recombine them in novel ways to generalize grammatical patterns. Although this does require buffer storage capacity for the individual units in formulation, it also depends heavily on the ability to process, manipulate, and recombine these units. Because WM measures both, it should serve as a good predictor of such abilities, more so than PSTM.

The manner in which participants were introduced to the grammatical structures and the way in which they got feedback on their productions may also account for some of the relationship between WM and grammar scores. Mackey (2006) has suggested that WM is particularly involved in learners' analysis of the input, especially that which gives feedback on error. Instead of being taught the grammatical morphemes and noun-adjective word order explicitly, as typically occurs in a foreign-language classroom setting, these participants were simply exposed to the structures in a sentence context. This more 'naturalistic' way of learning requires much more initiative and effort on the learners' part, and may also require a different type of language processing. Participants also did not get explicit feedback on their productions;

instead, they simply heard the correct utterance as feedback. To learn from this information, they had to maintain their own utterance in memory while actively comparing it to the feedback they heard. Such operations, involving both storage and processing, must be heavily dependent upon WM capacities.

### *Vocabulary and Grammar*

As hypothesized, a strong relationship was found between vocabulary and grammar scores with intercorrelations on the order of 0.44 to 0.76. Such strong correlations might be surprising if considered in the traditional L1 framework of a words-and-rules system, which posits that grammar and vocabulary are learned separately and rely on different processing mechanisms (Brown, 1973; Chomsky, 1957, 1965; Katz & Postal, 1964; Pinker, 1991). However, learners clearly cannot understand or induce grammatical patterns unless they can recognize the lexical components of utterances, and, reciprocally, there are clear contributions of grammatical understanding upon word learning (e.g., syntactic bootstrapping, Gleitman, 1990). Bates and Goodman (1997) and Marchman and Bates (1994) argue that in L1 acquisition, vocabulary and grammar are processed and learned by one unitary system, and that grammar necessarily depends upon the vocabulary it organizes. Bates and colleagues call this the critical mass hypothesis because it assumes vocabulary must reach a minimum critical mass size before grammar induction can occur.

This debate has taken place largely in the L1 literature, with few extensions made to L2 learning. Because of this paucity of discussion and relevant data, it is difficult to draw firm conclusions. Some dual-route models of L2 learning (Paradis, 1994; Ullman, 2001, 2005) nevertheless do allow for the possibility that vocabulary and grammar learning depend on the same mechanisms at the earliest stages of L2 learning. Above and beyond the differences



between L1 and L2, however, empirical research demonstrates the intimate interdependencies of lexis and grammar throughout language usage, both at the beginning and at later stages of learning. Corpus linguistics reveals many things about language: that much of communication makes use of fixed expressions memorized as formulaic chunks; that language is rich in collocational and colligational restrictions; that the phrase is the basic level of language representation where form and meaning meet with greatest reliability; that formulaic sequences play a central role in language acquisition; and that fluent language users have a vast repertoire of memorized language sequences (Pawley & Syder, 1983; Sinclair, 1991, 2004). Grammar and lexis are not entirely separable (Ellis, 2008; Römer, 2009). The recognition of their symbiosis lies at the core of modern developments in construction grammar (see Goldberg, 1995), cognitive linguistic (e.g., Ellis & Cadierno, 2009; Robinson & Ellis, 2008) and input-driven descriptions of language structure, processing, and acquisition (Collins & Ellis, 2009). Although the possibility of a L2 words-and-rules system is not ruled out by the results presented here, the strong interdependence of vocabulary and grammar, especially at the beginning stages of L2 learning, lead us more naturally to constructionist interpretations of the phenomena investigated.

#### *Vocabulary as a Mediating Factor*

Given the strong intercorrelations between grammar and lexis, the possibility of identifying any distinctions in their reliance on PSTM and WM was further investigated. Regression analyses were performed on the final grammar scores to determine whether their relationship with PSTM and WM was mediated by final vocabulary ability. The results showed that once final GS vocabulary scores were accounted for, neither PSTM nor WM contributed unique variance to GS grammar. Given the interactions of grammar and vocabulary in language structure and language processing, described previously, it is not surprising that final GS

vocabulary and grammar are as interrelated as this. Nevertheless, WM did contribute a unique 4% of variance to the final grammar scores overall (TP), which suggests that grammar induction did make more use of WM.

A clearer test of the mediation of vocabulary in the relationships between memory and grammar learning, as suggested by Service and Kohonen (1995), involved causal path analysis of vocabulary from the end of the training phase as a mediator between PSTM and WM and final GS vocabulary and grammar. This demonstrated substantial independent effects of PSTM (0.41) and WM (0.42) upon trained vocabulary, an effect of trained vocabulary upon GS grammar (0.33) that allowed for mediated indirect effects of PSTM (0.14) and WM (0.14) upon GS grammar, and direct effects of PSTM (0.25) and WM (0.30) upon GS grammar. Thus, there are significant independent effects of PSTM and WM upon L2 vocabulary learning and upon L2 grammar learning—some of which are mediated by vocabulary and some of which are direct effects. These memory systems are indeed involved in vocabulary learning, but they are also involved in grammar induction from language usage over and above that.

#### *Suggestions for Future Research*

As is usually the case, these results warrant replication in other populations and with other languages and language structures. Replicating the study with more complex artificial language systems, such as Brocanto (Friederici, Steinhauer, & Pfeifer, 2002; Opitz & Friederici, 2004) as well as real languages, would allow for converging evidence from various language systems. Although the current research has important implications for understanding language learning, there are limitations to how much the results from laboratory artificial language learning can be generalized to real-life language learning situations. Future research should also examine the longitudinal development of vocabulary and grammar abilities, rather than just their

final attainment and knowledge at one prior stage. Although a longitudinal design was used here, the current study lasted just two hours, thus limiting the number of time-points to chart development and the conclusions that could be drawn. The value of constructionist accounts of the codevelopment of grammar and lexis would also be informed by assessing phraseological and formulaic knowledge in development, and the degree to which these relied upon PSTM and WM. Varying the way in which the language is taught and in which feedback is given is also important to determine whether the relationships found in this study apply across learning conditions. Including grammaticality judgments may also be important to determine whether learners base what is possible based on occurrence versus nonoccurrence, or on more abstract knowledge.

One additional question for further research concerns the degree to which the PSTM-language learning correlations observed here are generalizable to other measures of these same skills but with different content. Nonword repetition tests are affected by lexical knowledge—thus Thorn and Gathercole (1999, 2001) showed that English-French bilinguals are equally good at repeating nonwords conforming to English and French phonotactics, but monolinguals are worse for nonwords conforming to the phonotactics of an unknown language. The PSTM items here and the words in the artificial L2 both conformed to L1-English phonemes and phoneme combinations, and some of the same nonwords were used as stimuli in both nonword repetition tasks and the artificial language. Phonological short-term memory tasks are better predictors of foreign language vocabulary learning when they are more word-like in the foreign language than the native language. One might expect, therefore, that experiments which use different L1 and foreign language phonotactics as well as PSTM content that is foreign-like, should maximize any relations between PSTM and language learning. This interesting empirical question awaits

further research.

Finally, a number of exploratory analyses that involved the explicit knowledge scores were presented. These analyses suggested strong relationships between explicit knowledge and final vocabulary and grammar scores. However, it is important to emphasize that these results must be interpreted cautiously. The scores were compiled from only two open-ended questions producing low outcome scores. Nevertheless, the analyses suggest that the relations between memory measures, vocabulary, and grammar abilities warrant further investigation in future studies that involve both explicit and implicit learning situations.

### Conclusions

Constructionist accounts naturally allow for a strong relationship between grammar and vocabulary while acknowledging their differential reliance upon PSTM and WM. Vocabulary learning involves the sequential sound patterns of words and their arbitrary mapping to meaning; grammar learning involves the sequential patterns of words and morphemes. Both involve the memorization of phonological sequences in PSTM, yet grammatical patterns are more global. They apply to the utterance as a whole, beyond the individual word, and more than that, they involve the abstraction of patterns over sets of morphemes. These more complicated patterns may demand more processing capacities, the holding of a greater amount of information over time, and the identification, selection, and correlation of relevant features both in the input and in long-term memory. Thus they are likely to be somewhat more reliant upon WM. Nevertheless, vocabulary and grammar are highly interrelated in use and learning, and both make substantial demands upon PSTM and WM.

## References

- Adams, A. M., & Gathercole, S. E. (1996). Phonological working memory and spoken language development in young children. *Quarterly Journal of Experimental Psychology*, *49*, 216-233.
- Adams, A. M., & Gathercole, S. E. (2000). Limitations in working memory: Implications for language development. *International Journal of Language & Communication Disorders*, *35*, 95-116.
- Arbuckle, J. L. (2006). Amos (Version 7.0) [Computer software]. Chicago: SPSS.
- Audacity Team. (2008). Audacity (Version 1.2.6) [Computer software]. Retrieved September 7, 2008 from <http://audacity.sourceforge.net/>
- Baddeley, A. D. (1998). Working memory. *C.R. Académie des sciences Paris, Sciences de la Vie / Life Sciences*, *321*, 167-173.
- Baddeley, A. D. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, *36*, 189-208.
- Baddeley, A. D., Gathercole, S. E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105*, 158-173.
- Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation: Vol. 8*. (pp. 47-90). San Diego, CA: Academic Press.
- Baddeley, A. D., Papagno, C., & Vallar, G. (1988). When long-term learning depends on short-term storage. *Journal of Memory and Language*, *27*, 586-595.
- Bates, E., & Goodman, J. C. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia and real-time processing. *Language and Cognitive Processes*, *12*, 507-584.

- Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language*, 44, 235-249.
- Brooks, P. J., Braine, M. D. S., Catalano, L., Brody, R. E., & Sudhalter, V. (1993). Acquisition of gender-like noun subclasses in an artificial language: The contribution of phonological markers to learning. *Journal of Memory and Language*, 32, 76-95.
- Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.
- Chomsky, N. (1957). *Syntactic structures*. Berlin: Mouton de Gruyter.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Collins, L. & Ellis, N. C. (Eds). (2009). Input and second language construction learning: Frequency, form, and function [Special issue]. *Modern Language Journal*, 93(3).
- Daneman M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450-466.
- Daneman, M., & Case, R. (1981). Syntactic form, semantic complexity, and short-term memory: Influences on children's acquisition of new linguistic structures. *Developmental Psychology*, 17, 367-378.
- DeKeyser, R. M. (2005). What makes learning second-language grammar difficult? A review of issues. *Language Learning*, 55, 1-25.
- Ellis, N. C. (1996). Sequencing in SLA: Phonological memory, chunking, and points of order. *Studies in Second Language Acquisition*, 18, 91-126.
- Ellis, N. C. (2008). Phraseology: The periphery and the heart of language. Preface to F. Meunier and S. Granger (Eds.), *Phraseology in language learning and teaching* (pp. 1-13). Amsterdam: Benjamins.

- Ellis, N. C. & Cadierno, T. (Eds). (2009). Constructing a second language: Introduction to the special section. *Annual Review of Cognitive Linguistics*, 7, 111-290.
- Ellis, N. C., & Sinclair, S. G. (1996). Working memory in the acquisition of vocabulary and syntax: Putting language in good order. *Quarterly Journal of Experimental Psychology*, 49, 234-250.
- Fan, M. Y. (2003). Frequency of use, perceived usefulness, and actual usefulness of second language vocabulary strategies: A study of Hong Kong learners. *The Modern Language Journal*, 87, 222-241.
- Fortkamp, M. B. M. (1999). Working memory capacity and aspects of L2 speech production. *Communication & Cognition*, 32, 259-296.
- French, L. M. (2006). *Phonological working memory and second language acquisition: A developmental study of Francophone children learning English in Quebec*. Lewiston, NY: Edwin Mellen Press.
- French, L. M., & O'Brien, I. (2008). Phonological memory and children's second language grammar learning. *Applied Psycholinguistics*, 29, 463-487.
- Friederici, A. D., Steinhauer, K., & Pfeifer, E. (2002). Brain signatures of artificial language processing. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 529-534.
- Gathercole, S. E., & Baddeley, A. D. (1993). *Working memory and language*. Mahwah, NJ: Erlbaum.
- Gathercole, S. E., Pickering, S. J., Hall, M., & Peaker, S. M. (2001). Dissociable lexical and phonological influences on serial recognition and serial recall. *Quarterly Journal of Experimental Psychology*, 54, 1-30.

- Gleitman, L. (1990). The structural source of verb meaning. *Language Acquisition, 1*, 3-55.
- Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.
- Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition, 14*, 25-38.
- Haywood, J. A., & Nahmad, H. M. (1965). *A new Arabic grammar of the written language*. London: Lund Humphries.
- Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*, 60-99.
- Jones, G., Gobet, F., & Pine, J. M. (2007). Linking working memory and long-term memory: A computational model of the learning of new words. *Developmental Science, 10*, 853-873.
- Jones, G., Gobet, F., & Pine, J. M. (2008). Computer simulations of developmental change: The contributions of working memory capacity and long-term knowledge. *Cognitive Science, 32*, 1148-1176.
- Katz, J. J., & Postal, P. M. (1964). *An integrated theory of linguistic descriptions*. Cambridge, MA: MIT Press.
- King, G. (2003). *Modern Welsh: A comprehensive grammar*. New York: Routledge.
- Kormos, J., & Sáfár, A. (2008). Phonological short-term memory, working memory and foreign language performance in intensive language learning. *Bilingualism: Language and Cognition, 11*, 261-271.



- Leeser, M. J. (2007). Learner-based factors in L2 reading comprehension and processing grammatical form: Topic familiarity and working memory. *Language Learning, 57*, 229-270.
- Mackey, A. (2006). From introspections, brain scans, and memory tests to the role of social context: Advancing research on interaction and learning. *Studies in Second Language Acquisition, 28*, 369-379.
- Mackey, A., Philp, J., Egi, T., Fujii, A., & Tatsumi, T. (2002). Individual differences in working memory, noticing of interactional feedback, and L2 development. In P. Robinson (Ed.), *Individual differences in instructed language learning* (pp. 181-209). Amsterdam: Benjamins.
- Marchman, V. A., & Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language, 21*, 339-366.
- Marton, K., & Schwartz, R. G. (2003). Working memory capacity and language processes in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 46*, 1138-1153.
- Masoura, E. V., & Gathercole, S. E. (1999). Phonological short-term memory and foreign language learning. *International Journal of Psychology, 34*, 383-388.
- O'Brien, I., Segalowitz, N., Collentine, J., & Freed, B. (2006). Phonological memory and lexical, narrative, and grammatical skills in second language oral production by adult learners. *Applied Psycholinguistics, 27*, 377-402.
- O'Brien, I., Segalowitz, N., Freed, B., & Collentine, J. (2007). Phonological memory predicts second language oral fluency gains in adults. *Studies in Second Language Acquisition, 29*, 557-582.

- Opitz, B., & Friederici, A. D. (2004). Brain correlates of language learning: The neuronal dissociation of rule-based versus similarity-based learning. *Journal of Neuroscience*, *24*, 8436-8440.
- Papagno, C., Valentine, T., & Baddeley, A. D. (1991). Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory and Language*, *30*, 331-347.
- Paradis, M. (1994). Neurolinguistic aspects of implicit and explicit memory: Implications for bilingualism and SLA. In N. C. Ellis (Ed.), *Implicit and explicit learning of languages*, (pp. 393-420). San Diego, CA: Academic Press.
- Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 191-225). London: Longman.
- Pinker, S. (1991). Rules of language. *Science*, *253*, 530-535.
- Polomé, E. (1967). *Swahili language handbook*. Washington DC: Center for Applied Linguistics.
- Robinson, P. (2002). Effects of individual differences in intelligence, aptitude, and working memory on adult incidental SLA: A replication and extension of Reber, Walkenfield, and Hernstadt (1991). In P. Robinson (Ed.), *Individual differences in instructed language learning* (pp. 181-209). Amsterdam: Benjamins.
- Robinson, P., & Ellis, N. C. (Eds.). (2008). *Handbook of cognitive linguistics and second language acquisition*. London: Routledge.
- Robinson, B. F., Mervis, C. B., & Robinson, B. W. (2003). The roles of verbal short-term memory and working memory in the acquisition of grammar by children with Williams syndrome. *Developmental Neuropsychology*, *23*, 13-31.

- Römer, U. (2009). The inseparability of lexis and grammar: Corpus linguistic perspectives. *Annual Review of Cognitive Linguistics*, 7, 140-162.
- Sanaoui, R. (1995). Adult learners' approaches to learning vocabulary in second languages. *Modern Language Journal*, 79, 15-28.
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime reference guide*. Pittsburgh, PA: Psychology Software Tools, Inc.
- Service, E. (1992). Phonology, working memory, and foreign-language learning. *Quarterly Journal of Experimental Psychology*, 45, 21-50.
- Service, E., & Kohonen, V. (1995). Is the relation between phonological memory and foreign language learning accounted for by vocabulary acquisition? *Applied Psycholinguistics*, 16, 155-172.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Sinclair, J. (2004). *Trust the text: Language, corpus and discourse*. London: Routledge.
- Speidel, G. E. (1993). Phonological short-term memory and individual differences in learning to speak: A bilingual case study. *First Language*, 13, 69-91.
- Sunderman, G., & Kroll, J. F. (2009). When study-abroad experience fails to deliver: The internal resources threshold effect. *Applied Psycholinguistics*, 30, 79-99.
- Thorn, A. S. C., & Gathercole, S. E. (1999). Language-specific knowledge and short-term memory in bilingual and non-bilingual children. *Quarterly Journal of Experimental Psychology*, 52, 303-324.
- Thorn, A. S. C., & Gathercole, S. E. (2001). Language differences in verbal short-term memory do not exclusively originate in the process of subvocal rehearsal. *Psychonomic Bulletin & Review*, 8, 357-364.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, 28, 127-154.

Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition*, 4, 105-122.

Ullman M. T. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In C. Sanz (Ed.), *Mind and context in adult second language acquisition: Methods, theory and practice* (pp. 141-178). Washington, DC: Georgetown University Press.

van den Noort, M. W. M. L., Bosch, P., & Hugdahl, K. (2006). Foreign language proficiency and working memory capacity. *European Psychologist*, 11, 289-296.

Williams, J. N., & Lovatt, P. (2003). Phonological memory and rule learning. *Language Learning*, 53, 67-121.

Williams, J. N. (2012). Working memory and SLA. In S. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition*. London: Routledge.

Author Note

Katherine I. Martin, Department of Psychology, University of Michigan; Nick C. Ellis, Department of Psychology and English Language Institute, University of Michigan. Katherine I. Martin is now at Department of Linguistics, University of Pittsburgh.

This work was supported by research grants from the University of Michigan Department of Psychology, Honors College, and Residential College, including the Muenzer Memorial Award, the Tanner Memorial Award, and the Otto Graf Scholarship. The authors would like to thank Andrew Mueller and Andrew Lorelli for assisting in the transcribing and scoring of participant utterances and Elizabeth Bogner and Ann Cassel for help in creating the illustrations for all the vocabulary words and sentence stimuli.

Correspondence regarding this work should be addressed to Katherine Martin, Department of Linguistics, 2832 Cathedral of Learning, University of Pittsburgh, Pittsburgh, PA, 15260. E-mail: kim20@pitt.edu Phone: (616) 402-4984.