Applied Psycholinguistics **27** (2006), 377–402
Printed in the United States of America
DOI: 10.1017.S0142716406060322

Phonological memory and lexical, narrative, and grammatical skills in second language oral production by adult learners

IRENA O'BRIEN
Université du Québec à Montréal and Centre for the Study of Learning and Performance

NORMAN SEGALOWITZ
Concordia University and Centre for the Study of Learning and Performance

JOE COLLENTINE
Northern Arizona University

BARBARA FREED
Carnegie Mellon University

Received: April 19, 2005 Accepted for publication: November 11, 2005

ADDRESS FOR CORRESPONDENCE
Irena O'Brien, Department of Psychology, McGill University, Montréal, Québec H3C 3J7, Canada.
E-mail: irena.obrien@gmail.com

ABSTRACT
This study investigated the role of phonological memory (PM) in second language (L2) speech production by English-speaking adults who were learning Spanish. PM, operationalized as serial nonword recognition, and L2 lexical, narrative, and grammatical abilities from speech samples were assessed 13 weeks apart. After controlling for the amount of speech output, PM contributed significantly to the development of L2 narrative skills for less proficient participants (17.5% of variance explained) and to gains in correct use of function words for more proficient participants (15.7% of variance explained). These findings suggest that PM plays an important role in narrative development at earlier stages of L2 learning and in the acquisition of grammatical competence at later stages.

The role of phonological memory (PM) in first language (L1) vocabulary learning has been well established (see Baddeley, Gathercole, & Papagno, 1998, for a review). For vocabulary acquisition to be successful, the learner needs to be able to hold and operate upon phonological sequences in working memory to process input correctly and thereby add vocabulary to his or her repertoire. There has been considerably less research, however, into the role of PM in the development of

© 2006 Cambridge University Press 0142-7164/06 $12.00

speech production, and to our knowledge, none that has looked at second language (L2) speech production in adults. Nevertheless, many of the same issues that link PM to vocabulary acquisition may also link it to the development of speech production abilities. In particular, it may be useful to investigate the role of PM in the development of grammatical abilities, as they manifest themselves in real-time speech. The purpose of this research was to investigate such a relationship in the L2 speech production of adults in the process of learning an L2.

By way of background, PM is a component of Baddeley and Hitch's (1974, and as revised by Baddeley, 1986, 2000) working memory model. This model contains a multifunctional system called the central executive and three subsidiary systems: the visuospatial sketchpad, the phonological loop (also called PM), and the episodic buffer. The central executive is responsible for complex processing operations, such as focusing, switching, and dividing attention, not attributed to the three subsidiary systems (Baddeley & Hitch, 1974; Baddeley, 1986, 2000). The visuospatial sketchpad is responsible for the short-term storage and processing of visual and spatial information. The PM component of the working memory system has two subcomponents: the phonological store is responsible for storing verbal information for short periods of time (~2 s), and is represented in a sound-based (phonological) code. The articulatory rehearsal system refreshes the contents of the phonological store so that information can be held in PM for longer than 2 s. PM capacity varies among individuals (Gathercole, Pickering, Hall, & Peaker, 2001; Karakas, Yalin, Irak, & Erzengin, 2002), increases throughout childhood and adolescence (Gathercole, 1998; Hulme, Thomson, Muir, & Lawrence, 1984; Karakas et al., 2002), and decreases slightly (but significantly) throughout adulthood (Karakas et al., 2002). The episodic buffer, a more recent addition to the model, combines information from long-term memory and the specialized subsidiary storage systems to create integrated episodes.

There is abundant evidence that PM is related to children's L1 vocabulary knowledge, even as late as 14 years of age, and to their ability to learn new words (see Baddeley et al., 1998, for a review). With regard to speech production, however, there have been only a few investigations into the role of PM in L1 speech. Studies examining L1 speech corpora in normally developing preschool children have found that children with better PM skills produce longer, more grammatically complex, and lexically richer utterances than children with less well-developed PM abilities (Adams & Gathercole, 1995, 2000). PM predicts the mean length of utterance better than does chronological age or mental age, and utterance length is correlated with sentence complexity such that children who produce longer utterances, and hence have better PM skills, show greater grammatical and semantic complexity (Blake, Austin, Cannon, Lisus, & Vaughan, 1994). Finally, PM predicts both the amount of story information recalled and the sentence length that children use in their narration (Adams & Gathercole, 1996). These studies suggest that PM is implicated in certain aspects of L1 speech production, namely in utterance length, and grammatical and semantic complexity.

Another line of evidence suggesting that PM may be related to L1 grammatical abilities comes from sentence repetition studies. Sentence repetition ability has been used as a measure of the acquisition of grammatical skill because it requires understanding relationships between words (Blake et al., 1994). Willis and

Gathercole (2001) found that children with better PM skills were more accurate at repeating complex, but not simple, sentences in their L1 than children with weaker PM abilities (see also Blake et al., 1994). The role of PM in L2 speech production has not been investigated directly. However, other L2 studies have shown that PM is correlated with children's L2 vocabulary (Masoura & Gathercole, 1999, 2005) and predicts the ability of both children (Cheung, 1996) and adults to learn new L2 vocabulary (Atkins & Baddeley, 1998; Gupta, 2003; Speciale, Ellis, & Bywater, 2004). With regard to L2 grammar acquisition, PM is implicated in the ability of children (Daneman & Case, 1981) and adults (Williams & Lovatt, 2003) to learn and generalize novel grammar learned in the laboratory. Ellis and Sinclair (1996), although they did not specifically measure PM, showed that requiring adults to maintain L2 utterances in PM through repetition (rehearsal) resulted in better production performance and metalinguistic knowledge about a morphological mutation rule compared to that of a group required to engage in articulatory suppression. French (2004a) found that, in elementary Francophone children enrolled in an intensive English program, PM measured at Time 1 correlated with specific tasks assessing knowledge of grammar at Time 2, but that this relationship was mediated by English vocabulary knowledge. In a new sample of participants, French (2004b) controlled for lexical content and again found that PM at Time 1 was significantly correlated with the grammar measures at Time 2. However, this association was not mediated by English vocabulary knowledge. Service (1992) and Service and Kohonen (1995) provide additional indirect evidence that PM may be related to L2 grammatical development. They showed that PM predicted Finnish elementary school children's scores on a variety of English tests, including reproduction of structures and written production (requiring some knowledge of English grammar). However, this relationship was also explained by English vocabulary knowledge. Finally, there are a number of case studies showing that PM is particularly involved in memory for function words and word order (Romani, 1994; Speidel, 1993).

From the review above, it appears that PM may be related to the development of grammatical skill. The research in this area, however, is far from conclusive, and suggests that the relationship between PM and grammatical abilities may be driven by vocabulary skills. Furthermore, very few of these studies have investigated speech production directly and, to our knowledge, none has explored the role of PM in L2 speech production by adults.

The present study investigates the association between PM and L2 speech production by adults. Specifically, the research reported here investigates the relationship between PM and the development of certain aspects of lexical, narrative, and grammar abilities obtained from speech samples collected from native English-speaking adults learning Spanish as an L2 over the course of a semester at university.

VARIABLES

PM was measured at the beginning (Time 1) and end (Time 2) of the semester using a serial nonword recognition task. This task required participants to judge whether the presentation order of two strings of nonwords was the same or different. The

task was chosen for several reasons. First, unlike serial nonword recall and nonword repetition (the two most frequently used measures of PM), serial nonword recognition is less susceptible to long-term memory effects, such as lexicality (Gathercole et al., 2001; Hulme, Maughan, & Brown, 1991), language of testing (Thorn, Gathercole, & Frankish, 2002), syllable frequency (Nimmo & Roodenrys, 2002), phonotactic knowledge (Gathercole & Pickering, 1999), and so forth. Second, serial nonword recognition may be a relatively pure measure of phonological storage uncontaminated by articulatory output demands (Baddeley, 2003). Gathercole and Pickering estimated that the basic capacity of the phonological store (which they called the intact parameter) could be measured in a serial recall task by using low probability nonwords. The phonotactic parameter, reflecting the contribution of phonotactic knowledge, was computed as the difference between the serial recall of high probability and low probability nonwords. The lexical parameter, reflecting the contribution of long-term lexical knowledge, was computed as the difference between the serial recall of words and high probability nonwords. They found that only the intact parameter (phonological store) best distinguished high from low vocabulary children, and that the serial recognition of both words and nonwords was highly correlated with the intact parameter, but not the phonotactic or lexical parameters.

The nonwords used in the serial nonword recognition task were taken from Gathercole et al. (2001) and followed English phonotactic rules (see Appendix A for a complete list of nonwords). English rather than Spanish nonwords were used to minimize the influence of the participants' familiarity with Spanish. All of the participants were native English speakers, and all were university students at English-speaking institutions. Therefore, it was reasonable to assume that they all possessed good English language skills. By using English nonwords, the participants were equated on the long-term knowledge of the phonotactic properties of the nonword stimuli used in the task. Thus, individual differences in performance on this task should reflect differences in the ability of participants to hold the items in PM rather than to differences in long-term phonological knowledge of Spanish.

Samples of Spanish speech production were obtained from recordings of participants' oral proficiency interviews (OPIs) at Time 1 and Time 2. Four-minute extracts from the interviews were selected at both times and analyzed for a number of measures: productive vocabulary, narrative abilities, accuracy (i.e., correct use) of 15 elements of inflectional morphology, and use of two complex grammatical structures (see Collentine, 2004, for details). Productive vocabulary was operationalized as the number of unique lexical items produced in each (Time 1 and Time 2) speech segment. Fourteen of the 15 elements of inflectional morphology (subjunctive accuracy was not retained, as it was seldom, if ever, used by the participants) were grouped into two broad categories: accuracy of free and bound morphemes. In grouping the elements of inflectional morphology into free and bound morphemes, cues were taken from the child development literature that, in constructing language, children tend to prefer free morphemes to bound morphemes (Slobin, 1985). The free morphemes were further grouped into function words and pronouns, based on evidence from single case studies that PM is important for remembering function words (Romani, 1994; Speidel, 1993). Subordinate clause use and coordinate clause use were the two grammatical structures that were retained (see the Method section for details).

Narrative abilities included five elements (past tense verbs, third person morphology, past participles, present participles, and public verbs, e.g., verbs of communication such as *admit, say*) identified by Biber (1988) in a factor analysis study of 481 written and spoken texts as occurring most frequently in narrative discourse. "Narrative abilities," as identified by Biber and used in the present study, is a lexicogrammatical measure such that when these features coexist in a given segment of discourse they have been found to be reliable predictors of narrative discourse. This measure is not meant to reflect grammatical competence but rather storytelling ability, and is thought to be consistent across languages. For example, Biber (1988) found that these same elements co-occurred in a sample of narratives in Somali. The following is a typical example of narrative discourse produced by one of the participants containing many events, past tense verbs, and public verbs. As reported in Collentine (2004, pp. 241–242), words in all capital letters reflect the student's omissions. Each break comprises a new turn:

> *En Granada el Alhambra fue muy bonito. Es un jardín que los moros construyeron y ahora es libre para la pública. Me gusta Granada. Es muy bonito . . . más que Alicante y fui a UN BAILE DE flamenco con los gitanos. Me gusta pero no fui a los bares y discotecas porque no me gusta y en Madrid fui a EL Prado y la reina Sofía y nada más porque nosotros estuvamos afuera DE Madrid en una ciudad muy pequeña y es muy caro para ir a Madrid. Es como de Estados Unidos a Nueva York pero no vi mucho porque no tengo tiempo y gastó mucho tiempo en el grado y cuando salí no camino sobre el ciudad porque no tengo tiempo pero no sé.*
>
> In Granada the Alhambra was very pretty. It's a garden that the Moors built and now it is free to the public. I like Granada. It's very pretty . . . more than Alicante and I went to A flamenco DANCE with the gypsies. I like it but I didn't go to the bars and discos because I don't like {that/them} and in Madrid I went to THE Prado and Queen Sofia and nothing else because we were outside OF Madrid in a small city and it is expensive to go to Madrid. It's like from the United States to New York because I didn't see much because I do not have time and I spend/spent a lot of time on the grade [*sic:* on my grades] and when I left I do not walk over the city because I do not have time but I don't know.

As mentioned above, there is some evidence that the relationship between PM and grammatical development is mediated by vocabulary growth. Consequently, people may use more grammatical elements simply because they produce more words. Thus, it is possible that those who gained in oral competence between Time 1 and Time 2 may merely have used more words at Time 2 than at Time 1, thereby obscuring a true gain. In order not to overstate the role of PM in such oral gains, and to assess the amount of oral gain that was due to an increase in the total number of words produced, the gain in the total number of words spoken in the extracts was entered into the analyses.

It is also possible that the gain in speech production measures may have been at least partly due to the number of courses in Spanish in which the participants were enrolled (between one and five). Consequently, again in order to not overstate the role of PM, the number of courses was entered into the analyses.

There is evidence that in children PM is more important at earlier stages of L1 (Gathercole, Willis, Emslie, & Baddeley, 1992) and L2 learning (Cheung,

1996; French, 2004a). More recently, O'Brien, Segalowitz, Freed, and Collentine (2006) extended this finding to adults, demonstrating that PM was more strongly associated with L2 oral fluency development in the less proficient adult L2 learners than in the more proficient learners. We wished to look at whether this pattern of results would extend to other measures of adult L2 speech production. Specifically, we wished to see whether the relationship between PM and Spanish speech production, as measured by productive vocabulary, narrative abilities, and grammatical skill, was different for the less proficient participants compared to the more proficient participants. The SAT II Spanish Test (The College Board, 2000), a test of reading ability in Spanish that measures knowledge of vocabulary, structure, and idiomatic expressions, was used to assess participants' levels of Spanish competence on entry into the study and to distinguish low ability from high ability learners of Spanish.

In summary, the present study asked two questions. Is PM implicated in the development of specific aspects of L2 speech production, namely, productive vocabulary, narrative abilities, and specific grammatical abilities? Is the relationship between PM and the above measures stronger for participants with initially lower levels of ability than for those with initially higher levels of ability? It was hypothesized that PM would be positively related to the development of L2 speech production (i.e., productive vocabulary, narrative abilities, and specific grammatical abilities) in the earlier phases of language learning, that is, in the initially less proficient participants as opposed to the initially more proficient participants.

METHOD

Participants

Participants were 43 ($M = 21.84$ years of age, $SD = 7.00$; 33 females) native speakers of English studying Spanish as an L2. The participants all had at least two prior semesters of formal study of Spanish; they had not previously studied Spanish abroad, Spanish was not a heritage language, and Spanish was not spoken at home. Their SAT II Spanish scores at the start of the semester varied between 310 (equivalent to first year of study) and 720 (equivalent to third year of study; $M = 470, SD = 101$). Eighteen participants were enrolled in one Spanish course at the University of Colorado in the United States and the remaining 25 participants, also from the United States, spent a semester in Spain at the Universidad de Alicante in a study abroad program and took between three and five courses.

Procedure

The students were tested on PM and participated in a 20- to 30-min, tape-recorded OPI in Spanish at the beginning of the semester (Time 1). At this time, to assess their initial level of Spanish knowledge, they completed the SAT II Spanish Test (excluding the listening portion). At the end of the semester (Time 2), they were tested again on PM and participated in another OPI.

PM. Participants were tested on the serial order recognition of lists of nonwords at both Time 1 and Time 2. The test of serial recognition required participants to judge

whether the second presentation of a list of items was the same (same trials) or different (different trials) from the first presentation. The stimulus set consisted of one-syllable consonant–vowel–consonant nonwords taken from Gathercole et al. (2001). There were eight lists of nonwords at each of three list lengths: five, six, and seven. The lists were constructed in such a manner that, within each list length, all vowel sounds were different and consonants were as phonologically distinct as possible. Within each list, the items were ordered randomly across participants. Four of the lists at each list length were same trials (the second presentation was identical to the first presentation) and four were different trials. On different trials, although the list items were identical, in the second presentation, one pair of items was transposed, with the restraint that the first and last pair of items were never transposed. This restraint was included to encourage the participants to process the complete stimulus list by reducing the salience of transposed items. The location of the transposed pair was varied randomly across list lengths and participants. A short pause was inserted between the first and second presentation of each list. A four-item practice list was also constructed, consisting of two same and two different trials. The following is an example of a same trial "charn, nig, kom, jeel, gadge, lerb, mun . . . charn, nig, kom, jeel, gadge, lerb, mun" and a different trial "chig, nam, peb, gop, jooch, lart, teed . . . chig, nam, gop, peb, jooch, lart, teed."

Participants were tested individually starting with the practice trials. If they made an error on the practice trials, the task was explained again and they proceeded to do the practice trials a second time. After completing the practice trials without error, participants were presented with the eight five-item lists, followed by the eight six-item lists, and then the eight seven-item lists. The nonwords were presented auditorily through earphones on a Macintosh computer using the built-in "Victoria, high quality" computer-generated voice that read the nonwords from a text file at the rate of approximately one item every 750 ms. There was a delay of approximately 1.5 s between the first and second presentation of each list that made up a given trial. After the second presentation of each list on a given trial, the participants indicated by pressing a key whether both lists in the trials were in the same or in a different order. The total number of correct responses was recorded for each participant.

OPI. At both Time 1 and Time 2, each student took part in a 20- to 30-min OPI in Spanish given by testers trained by the American Council of Teachers of Foreign Languages (Breiner-Sanders, Lowe, Miles, & Swender, 2000).[1] The interviews were tape-recorded and two 2-min extracts from each interview were retained, the first extract beginning at approximately 7 min and the second extract beginning at approximately 12 min. This yielded a 4-min Time 1 oral extract and a 4-min Time 2 oral extract for each student (see Segalowitz & Freed, 2004, for full details). The extracts were analyzed for the following Time 1 and Time 2 speech production measures.

Productive vocabulary. Each unique word produced by the participant in the extract at each test time was recorded and summed to yield a Time 1 and a Time 2 productive vocabulary score. Neologisms such as *problemo* and *overlapo*, representing only 0.6% of the total words in the corpus, were scored as unique

words. This score reflected the number of lexical items accessed by the participant at each test time.

Narrative abilities. The number of occurrences of the five elements identified by Biber (1988) as being indicative of narrative discourse was recorded: past tense verbs, third person morphology, past participles, present participles, and public verbs (i.e., verbs of communication such as *admit* and *say*). The sum across these occurrences at each test time yielded a Time 1 and a Time 2 narrative score. Occurrences of the Spanish imperfect and the preterit were included in this measure. Consistent with the finding that L2 learners do not substitute marked forms where unmarked forms are required (Lafford & Collentine, 1987, 1989), participants were credited with an item whether or not it was used correctly. An examination of the corpus showed that substitutions that employed a past tense form when a present tense form would have been more appropriate comprised less than 1% of the past tense instances.

Inflectional morphology. The number of accurate uses of 14 elements of inflectional morphology at each test time was recorded (subjunctive accuracy was not retained in this study because it was rarely, if ever, produced by participants; see Collentine, 2004, for details). These elements were then standardized across Times 1 and 2 as follows. One composite z score was computed across Times 1 and 2 for each element resulting in a combined (Time 1 + Time 2) mean score of zero and separate Time 1 and Time 2 z scores that were different from zero. In this manner, the relationship between Time 1 and Time 2 scores was maintained.

The z scores were then grouped into *free* and *bound* grammatical morphemes. The measure of *free grammatical morphemes* at Time 1 consisted of the mean of Time 1 z scores for accurate use of prepositions, object pronouns, coordinate conjunctions, subordinate conjunctions, plural pronouns, and feminine pronouns. The measure of Time 2 free grammatical morphemes reflected the mean of the Time 2 z scores for the same items. Similarly, measures of Time 1 and Time 2 *bound grammatical morpheme* use were calculated using the mean of Time 1 and Time 2 z scores, respectively, for correct use of copulas, present tense verbs, past tense verbs, indicatives, person, plural adjectives, plural verbs, and feminine adjectives.

Because PM seems to be especially important for remembering function words (Romani, 1994; Speidel, 1993), the free grammatical morphemes were further divided into *function words* (accurate use of prepositions, and coordinate and subordinate conjunctions) and *pronouns* (accurate use of object pronouns, plural pronouns, and feminine pronouns).

Clauses. Time 1 and Time 2 measures of *subordinate* and *coordinate clause use* were obtained based on the frequency with which these structures occurred.

Total words. Because participants produced more words in the Time 2 extracts than in the Time 1 extracts, it was conceivable that a gain in the speech production measures could be due to an increase in the number of words that participants produced. To control for this possibility, the gain in the total number of words

spoken between Time 1 and Time 2 was entered into the analyses. The gain was computed by residualizing total words produced at Time 2 against total words produced at Time 1, yielding a residualized gain score. This residualized gain score reflected participants' improvement in the total number of words they produced between Time 1 and Time 2, after equating them on Time 1 for total words produced.

SAT II Spanish Test. The SAT II Spanish test (The College Board, 2000) is used to place high school students into universities in the United States. For purposes of the present study, this test was used to evaluate the participants' overall Spanish language knowledge on entry into the study. The test contains 85 questions divided into three parts, each weighted equally: vocabulary and structure, paragraph completion, and reading comprehension. For this test, the students' raw scores are scaled on a range of 200–800 according to a linear function derived from norming techniques. The US mean (i.e., those around the 50th percentile) scaled score is 596 ($N = 25{,}155$, $SD = 30$). The means obtained in the present study are reported in Table 1.

RESULTS

This study asked whether PM was implicated in the development of specific measures of L2 speech production, namely productive vocabulary, narrative abilities, and specific grammatical abilities. It further asked whether this relationship was greater for the participants with initially lower levels of L2 ability than for those with initially higher levels of L2 ability.

The data were first inspected for outliers. For each variable, scores greater than 2 *SD* from the group mean for that variable were assigned the next highest score plus one. Scores more than 2 *SD* below the mean were assigned the next lowest score minus one. Unless otherwise indicated, the alpha level selected for significance was .05 (two tailed).

Entire cohort

The first question addressed was whether, over the entire cohort, PM was implicated in the development of L2 productive vocabulary, narrative abilities, and grammatical competence. Means and standard errors of all 10 variables (SAT II Spanish, total words, number of Spanish courses, PM, and six oral measures) for Time 1 and Time 2 are presented in Table 1, except for SAT II Spanish and number of Spanish courses, for which only a Time 1 value is available. To test whether the participants improved on PM and Spanish oral measures between Time 1 and Time 2, a series of a priori paired t tests was performed comparing performance on the measures at Time 1 and Time 2. The results of the t tests showed that, between Time 1 and Time 2, the participants gained in total words, $t\ (42) = 5.26$, $p < .001$, productive vocabulary, $t\ (42) = 5.98$, $p < .001$, correct use of free grammatical morphemes, $t\ (42) = 4.57$, $p < .001$, correct use of bound grammatical morphemes, $t\ (42) = 7.92$, $p < .001$, subordinate clause use,

Table 1. *Means (standard errors) for SAT II Spanish, total words, number of Spanish courses, phonological memory, and oral production measures at Time 1 and Time 2*

Variable	Entire Cohort ($n = 43$)		Low Ability ($n = 23$)		High Ability ($n = 20$)	
	Time 1	Time 2	Time 1	Time 2	Time 1	Time 2
SAT II Spanish	470 (15)	—	399 (11)	—	552 (18)	—
Total words	171 (6.51)	216 (9.60)***	174 (9.73)	220 (14.80)**	169 (0.37)	212 (12.00)**
No. of Spanish courses	2.63 (0.23)	—	2.74 (0.30)	—	2.50 (0.37)	—
Phonological memory	15.53 (0.47)	15.53 (0.38)	15.00 (0.63)	15.57 (0.56)	16.15 (0.69)	15.50 (0.53)
Productive vocabulary	79.00 (2.25)	99.88 (3.93)***	78.65 (2.99)	100.61 (5.79)***	79.40 (3.50)	99.05 (5.34)***
Narrative abilities	19.86 (1.00)	21.19 (1.31)	18.83 (1.37)	21.87 (1.98)	21.05 (1.46)	20.40 (1.71)
Grammatical morphemes[a]						
Free	−0.200 (0.05)	0.187 (0.08)***	−0.261 (0.07)	0.193 (0.13)**	−0.131 (0.09)	0.180 (0.11)*
Bound	−0.309 (0.07)	0.265 (0.09)***	−0.325 (0.10)	0.194 (0.13)***	−0.291 (0.10)	0.347 (0.13)***
Subordinate clause use	3.65 (0.45)	4.98 (0.46)*	2.96 (0.55)	4.87 (0.65)*	4.45 (0.69)	5.10 (0.67)
Coordinate clause use	5.58 (0.37)	6.74 (0.43)*	5.74 (0.49)	6.91 (0.61)	5.40 (0.56)	6.55 (0.62)*

[a]Free grammatical morphemes and bound grammatical morphemes are expressed as z scores.
*$p < .05$. **$p < .01$. ***$p < .001$.

t (42) = 2.57, p = .014, and coordinate clause use, t (42) = 2.51, p = .016. The participants did not improve in PM, t (42) = 0.00, *ns*, or narrative abilities, t (42) = 0.97, *ns*, between Time 1 and Time 2. Given that PM did not change over time, the remaining analyses use PM at Time 1.

Table 2 shows the Pearson intercorrelation coefficients for PM and the six Spanish oral measures across Times 1 and 2, except for PM for which only the Time 1 measure is reported (SAT II Spanish, total words, and number of courses were not included in this analysis). Time 1 PM was significantly correlated with Time 1 and Time 2 productive vocabulary (Time 1: p = .021; Time 2: p = .048), narrative abilities (Time 1: p = .022; Time 2: p = .043), free grammatical morphemes (Time 1: p = .049; Time 2: p = .008), and subordinate clause use (Time 1: p = .006; Time 2: p = .001). PM did not correlate with bound grammatical morphemes or coordinate clause use at either Time 1 or Time 2.

To test whether PM was related to the development of oral abilities between Time 1 and Time 2, residualized change scores were computed by regressing Time 2 oral performance on Time 1 oral performance. The resulting oral development measures reflect the amount that participants' scores changed between Time 1 and Time 2, equated on Time 1 scores. The Pearson correlation coefficients among the residualized change scores and PM are presented in Table 3. PM was significantly correlated with three out of the six oral development measures: narrative abilities, free grammatical morphemes, and subordinate clause use (all ps < .014). PM did not correlate with change in productive vocabulary, bound grammatical morphemes, or coordinate clause use.

To summarize, the above analyses revealed that, over the entire cohort, PM was related to development of narrative abilities, correct use of free morphemes, and subordinate clause use, such that those with better PM skills made more gains on these measures. Although PM was related to productive vocabulary at both Times 1 and 2, it did not predict change in vocabulary production over the semester.

The above analyses show that PM is implicated in a number of oral development measures. However, it is possible that the oral development was due, at least in part, to the greater number of words that participants used at Time 2 compared to Time 1 (i.e., participants who produced more speech may have used proportionately more lexical, narrative, and grammatical elements). To test whether the gains in oral measures may have been due to the increase in the total number of words used between Times 1 and 2, a series of hierarchical regressions was performed. A residualized change score was computed for total words used (Time 2 residualized against Time 1) and entered in Step 1 of the regression. We also wished to see whether in-class Spanish contact contributed to oral development. Consequently, the number of Spanish courses that participants' were enrolled in was entered in Step 2. So as not to overstate its effect, PM was entered in the last step, after the other control variables had been entered. At this point, given the evidence that PM may be important for remembering function words (Romani, 1994; Speidel, 1993), it was decided to divide the free grammatical morphemes into function words (containing prepositions, and subordinate and coordinate conjunctions) and pronouns. Separate hierarchical regressions were also performed on these two variables.

Table 2. *Simple intercorrelations among phonological memory and oral production measures at Time 1 and Time 2*

Entire Cohort ($n = 43$)													
Variable	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Phonological memory (1)	—												
2. Vocabulary (1)	.351*	—											
3. Narrative abilities (1)	.348*	.446**	—										
4. Free morphemes (1)	.303*	.607***	.660***	—									
5. Bound morphemes (1)	.074	.543***	.599***	.492**	—								
6. Subordinate clauses (1)	.416**	.620***	.433**	.516***	.427**	—							
7. Coordinate clauses (1)	−.030	.534***	.568***	.541***	.544**	.300*	—						
8. Vocabulary (2)	.304*	.468**	.262	.227	.273	.318*	.107	—					
9. Narrative abilities (2)	.310*	.220	.319*	.231	.328*	.075	.269	.596***	—				
10. Free morphemes (2)	.402**	.430**	.284	.311*	.372*	.254	.249	.779***	.760***	—			
11. Bound morphemes (2)	.241	.432**	.442**	.410**	.628***	.290	.377*	.679***	.665***	.804***	—		
12. Subordinate clauses (2)	.497**	.523***	.461**	.386*	.353*	.358*	.335*	.590***	.552***	.571***	.503***	—	
13. Coordinate clauses (2)	.125	.287	.302*	.320*	.384*	.132	.337*	.699***	.612***	.809***	.762***	.513***	—
Low Ability ($n = 23$)													
1. Phonological memory (1)	—												
2. Vocabulary (1)	.296	—											
3. Narrative abilities (1)	.016	.534**	—										
4. Free morphemes (1)	.278	.610***	.740***	—									
5. Bound morphemes (1)	−.012	.662***	.701***	.651***	—								
6. Subordinate clauses (1)	.381	.639***	.384	.655***	.465*	—							
7. Coordinate clauses (1)	−.025	.626**	.751***	.510*	.586**	.371	—						
8. Vocabulary (2)	.423*	.489*	.385	.351	.446*	.124	.254	—					

9. Narrative abilities (2)	.435*	.157	.268	.327	.271	.029	.100	.781***	—				
10. Free morphemes (2)	.483*	.376	.214	.296	.351	.139	.100	.855***	.831***	—			
11. Bound morphemes (2)	.238	.450*	.477*	.445*	.765***	.219	.264	.787***	.705***	.748***	—		
12. Subordinate clauses (2)	.471*	.367	.348	.320	.325	.187	.310	.693***	.677***	.682***	.643***	—	
13. Coordinate clauses (2)	.291	.356	.481*	.488*	.480*	.164	.259	.714***	.797***	.832***	.766***	.732***	—

High Ability ($n = 20$)

1. Phonological memory (1)	—												
2. Vocabulary (1)	.411	—											
3. Narrative abilities (1)	.684**	.359	—										
4. Free morphemes (1)	.279	.616**	.566**	—									
5. Bound morphemes (1)	.165	.419	.478*	.342	—								
6. Subordinate clauses (1)	.394	.633**	.437	.362	.403	—							
7. Coordinate clauses (1)	−.008	.449*	.413	.618**	.482*	.294	—						
8. Vocabulary (2)	.175	.453*	.119	.116	.028	.591**	−.083	—					
9. Narrative abilities (2)	.194	.317	.450*	.175	.431	.194	.496*	.286	—				
10. Free morphemes (2)	.317	.515*	.407	.361	.410	.435	.459*	.600**	.691**	—			
11. Bound morphemes (2)	.206	.415	.374	.352	.437*	.320	.544*	.523*	.695**	.918***	—		
12. Subordinate clauses (2)	.532*	.695**	.599**	.457*	.385	.542*	.372	.361	.480*	.411	.330	—	
13. Coordinate clauses (2)	−.041	.215	.118	.192	.267	.147	.421	.467*	.552*	.784***	.793***	.245	—

Note: The numbers after the variables denote testing time.
*$p < .05$. **$p < .01$. ***$p < .001$.

Table 3. *Zero-order intercorrelations among phonological memory and oral production residualized change scores*

Variable	1	2	3	4	5	6	7
Entire Cohort ($n = 43$)							
1. Phonological memory	—						
2. Vocabulary	.158	—					
3. Narrative abilities	.457**	.415**	—				
4. Free morphemes	.324*	.686***	.604***	—			
5. Bound morphemes	.251	.656***	.499**	.729***	—		
6. Subordinate clauses	.372*	.361*	.345*	.470**	.377*	—	
7. Coordinate clauses	.144	.632***	.453**	.762***	.690***	.443**	—
Low Ability ($n = 23$)							
1. Phonological memory	—						
2. Vocabulary	.3410	—					
3. Narrative abilities	.651***	.612**	—				
4. Free morphemes	.432*	.804***	.797***	—			
5. Bound morphemes	.374	.854***	.690***	.823***	—		
6. Subordinate clauses	.355	.674***	.518**	.644***	.706***	—	
7. Coordinate clauses	.309	.670***	.657**	.776***	.778***	.646***	—
High Ability ($n = 20$)							
1. Phonological memory	—						
2. Vocabulary	−.050	—					
3. Narrative abilities	.192	.124	—				
4. Free morphemes	.225	.487*	.293	—			
5. Bound morphemes	.115	.519*	.323	.723***	—		
6. Subordinate clauses	.445*	−.141	.079	.153	.083	—	
7. Coordinate clauses	−.042	.575*	.156	.743***	.661**	.123	—

*$p < .05$. **$p < .01$. ***$p < .001$.

The results of the hierarchical regressions (see Table 4) showed that, although changes in the total number of words that were used accounted for a large, significant amount of variance for all oral development measures (ΔR^2 between 13 and 77%), after Spanish courses were partialed out, PM still accounted for a further significant amount of variance for two variables: narrative abilities and subordinate clause use. With respect to free grammatical morphemes, after total words and number of Spanish courses were partialed out, the variance accounted for by PM showed a trend toward significance. This variance was accounted for by function words. PM did not account for a significant amount of variance for productive vocabulary, bound grammatical morphemes, coordinate clause use, or pronouns.

In summary, although change in total words used accounted for large amounts of variance in the oral development scores, PM still accounted for significant amounts of variance in narrative abilities and subordinate clause use, and showed a trend toward significance in free grammatical morphemes, which was explained by change in the correct use of function words.

Table 4. *Summary of hierarchical multiple regression analyses for variables predicting Spanish oral development for the entire cohort*

Variable	R	R^2	Adj. R^2	ΔR^2	ΔF	df	p
Δ Narrative abilities							
Δ Total words	.569	.323	.307	.323	19.595	1, 41	.000
No. of Spanish courses	.610	.372	.341	.049	3.091	1, 40	.086
Phonological memory	.666	.484	.444	.112	8.455	1, 39	.006
Δ Vocabulary							
Δ Total words	.880	.774	.768	.774	140.397	1, 41	.000
No. of Spanish courses	.880	.775	.764	.001	0.141	1, 40	.710
Phonological memory	.880	.775	.758	.000	0.061	1, 39	.806
Δ Free grammatical morphemes							
Δ Total words	.761	.580	.569	.580	56.517	1, 41	.000
No. of Spanish courses	.774	.598	.578	.019	1.881	1, 40	.178
Phonological memory	.791	.626	.597	.028	2.870	1, 39	.098
Δ Bound grammatical morphemes							
Δ Total words	.699	.488	.476	.488	39.139	1, 41	.000
No. of Spanish courses	.699	.489	.463	.000	0.028	1, 40	.867
Phonological memory	.709	.502	.464	.014	1.068	1, 39	.308
Δ Subordinate clause use							
Δ Total words	.402	.161	.141	.161	7.855	1, 41	.008
No. of Spanish courses	.403	.162	.120	.001	0.037	1, 40	.848
Phonological memory	.501	.251	.193	.089	4.606	1, 39	.038
Δ Coordinate clause use							
Δ Total words	.731	.534	.522	.534	46.922	1, 41	.000
No. of Spanish courses	.732	.536	.512	.002	0.157	1, 40	.694
Phonological memory	.732	.536	.500	.000	0.001	1, 39	.979
Δ Function words							
Δ Total words	.824	.679	.671	.679	86.776	1, 41	.000
No. of Spanish courses	.834	.695	.680	.016	2.134	1, 40	.152
Phonological memory	.866	.749	.730	.054	8.422	1, 39	.006
Δ Pronouns							
Δ Total words	.363	.131	.110	.131	6.206	1, 41	.017
No. of Spanish courses	.374	.140	.097	.009	0.409	1, 40	.526
Phonological memory	.393	.155	.090	.015	0.670	1, 39	.418

Low ability and high ability participants

The second question addressed was whether the relationship between PM and oral production measures observed over the entire cohort was greater for participants with initially lower levels of L2 ability than for those with initially higher levels of L2 ability. Previously reported research has suggested that PM is implicated in language learning for those at earlier rather than later stages of learning (Cheung, 1996; French, 2004a). To test for differences in the role of PM in learners at earlier rather than later stages of L2 learning, the entire cohort was divided into low and high ability participants based on their Spanish SAT II scaled scores. A median split could only be approximated, with 23 participants included in the low ability group (<460; $M = 20.43$ years of age, $SD = 1.50$; 18 females) and 20

participants included in the high ability group (≥460; $M = 23.45$ years of age, $SD = 10.02$; 15 females). Means and standard errors of all 10 variables (SAT II Spanish, total words, number of Spanish courses, PM, and six oral measures) at Times 1 and 2 for the low and high ability participants are shown in Table 1, except for SAT II Spanish and number of Spanish courses for which only a Time 1 value is available. A priori t tests used to test for differences between the groups at Time 1 revealed that on entry into the study, as expected given the selection procedure, the low ability participants had lower SAT II Spanish scores than the high ability participants, low ability: $M = 399$, $SE = 11$; high ability: $M = 552$, $SE = 18$; $t\ (41) = 7.63$, $p < .001$. The groups did not differ on total words, the number of Spanish courses they took, or on oral measures. However, the high ability participants showed a trend toward greater subordinate clause use ($p = .095$) than low ability participants.

The eight variables (PM, total words, and six oral measures) measured at Time 1 and Time 2 were then submitted to separate 2 × 2 mixed analyses of variance with time (1, 2) as the within factor and group (low ability, high ability) as the between factor. There was a main effect of time, but no main effect of group and no interaction effects, for total words, $F\ (1, 41) = 26.70$, $MSE = 1610.15$, $p < .001$, partial $\eta^2 = .394$; productive vocabulary, $F\ (1, 41) = 34.52$, $MSE = 268.24$, $p < .001$, partial $\eta^2 = .457$; free grammatical morphemes, $F\ (1, 41) = 20.11$, $MSE = 0.16$, $p < .001$, partial $\eta^2 = .329$; bound grammatical morphemes, $F\ (1, 41) = 62.79$, $MSE = 0.11$, $p < .001$, partial $\eta^2 = .605$; subordinate clause use, $F\ (1, 41) = 6.23$, $MSE = 5.64$, $p = .017$, partial $\eta^2 = .132$; and coordinate clause use, $F\ (1, 41) = 6.11$, $MSE = 4.73$, $p = .018$, partial $\eta^2 = .130$, indicating that both groups gained on these measures between Time 1 and Time 2. There were no significant main or interaction effects for PM or narrative abilities, indicating no difference on these measures between Time 1 and Time 2 or between groups.

Pearson intercorrelations among PM and oral measures at Times 1 and 2 are shown in Table 2, and among PM and residualized change scores, in Table 3, separately for the low and high ability groups. For the low ability group, PM did not predict any Time 1 measures, but predicted Time 2 productive vocabulary ($p = .04$), narrative abilities ($p = .04$), correct use of free morphemes ($p = .02$), and subordinate clause use ($p = .02$; see Table 2). For this group also, PM predicted change in narrative abilities ($p = .001$), and correct use of free morphemes ($p = .04$; see Table 3). For the high ability group, PM predicted narrative abilities at Time 1 ($p = .001$), and subordinate clause use at Time 2 ($p = .02$; see Table 2), together with change in subordinate clause use ($p = .05$; see Table 3).

To summarize, for the low ability group, PM was correlated with development of narrative abilities and correct use of free grammatical morphemes, such that those with better PM abilities made more gains in these measures. For this group also, although PM was significantly correlated with productive vocabulary and subordinate clause use at Time 2, it did not predict gains in these measures. For the high ability participants, PM predicted subordinate clause use at Time 2 together with gains in its use, such that those with better PM abilities produced proportionately more subordinate clauses at Time 2.

Separate hierarchical regression analyses were then computed for each ability group (low, high) and for each residualized oral measure (see above): productive vocabulary, narrative abilities, correct use of free grammatical morphemes, correct

use of bound grammatical morphemes, subordinate clause use, coordinate clause use, correct use of function words, and correct use of pronouns (see Table 5).

For the low ability group, change in the total number of words used accounted for significant, large amounts of variance for all the oral measures (ΔR^2 between .24 and .91). After total words were partialed out, the number of Spanish courses accounted for a further significant amount of variance for subordinate clause use only ($\Delta R^2 = .15$). PM explained a final significant amount of variance for narrative abilities only ($\Delta R^2 = .18$). For the high ability group, once total words (ΔR^2 between .04, *ns*, and .74), and number of Spanish courses (all ΔR^2 *ns*) were partialed out, PM explained a significant amount of variance only for accurate use of function words ($\Delta R^2 = .16$). The variance explained by PM for subordinate clause use showed a trend toward significance ($\Delta R^2 = .14$).

In summary, the low and high ability groups showed different patterns of results. For the low ability participants, PM was a significant predictor only of gains in narrative abilities after controlling for change in total words and number of Spanish courses. For the high ability participants, after change in total words and number of Spanish courses were partialed out, PM predicted gains in the correct use of function words. The relationship between PM and change in subordinate clause use showed a trend toward significance for this group. The groups also showed a different pattern of results with respect to change in total words. For the low ability participants, total words accounted for a large, significant amount of variance for all oral development measures. For the high ability participants, total words accounted for a smaller, significant amount of variance for only five oral development measures.

DISCUSSION

This study investigated the role of PM in adult L2 speech production. Specifically, it was asked whether, over the course of a semester, PM was implicated in gains made by adults studying Spanish as a L2 in productive vocabulary, narrative abilities, and a number of grammatical abilities. It was further asked whether these relationships would be different for participants who initially were categorized as low versus high ability in the language.

PM and productive vocabulary

The results indicated that, over the entire cohort, PM was related to vocabulary *use*. Specifically, those with better PM abilities produced a larger repertoire of words at both test times than those with weaker PM skills. The correlations reported in the present study are comparable to the correlations between PM and vocabulary use obtained in analyses of children's L1 speech samples (Adams & Gathercole, 1995, $r = .41$; Adams & Gathercole, 2000, rs $= .48$–$.62$).

In the study reported here, whereas PM was related overall to vocabulary use, it was not related to the *gain* in unique vocabulary items produced for the entire cohort or for the low and high ability groups. This result was somewhat unexpected, given that PM has been shown to predict the rate at which children acquire L1 (Gathercole et al., 1992) and L2 vocabulary (French, 2004a). Our results may have been obtained for several reasons. First, the vocabulary measure used

Table 5. *Summary of hierarchical multiple regression analyses for variables predicting Spanish oral development for the low ability and high ability groups*

Variable	R	R^2	Adj. R^2	ΔR^2	ΔF	df	p
Low Ability ($n = 23$)							
Δ Narrative abilities							
Δ Total words	.730	.532	.510	.532	23.902	1, 21	.000
No. of Spanish courses	.739	.547	.502	.015	0.641	1, 20	.433
Phonological memory	.850	.722	.678	.175	11.199	1, 19	.003
Δ Vocabulary							
Δ Total words	.895	.801	.792	.801	84.789	1, 21	.000
No. of Spanish courses	.898	.806	.786	.004	0.454	1, 20	.508
Phonological memory	.898	.806	.776	.000	0.043	1, 19	.839
Δ Free grammatical morphemes							
Δ Total words	.866	.749	.737	.749	62.716	1, 21	.000
No. of Spanish courses	.872	.761	.737	.012	1.008	1, 20	.327
Phonological memory	.880	.775	.739	.014	1.166	1, 19	.294
Δ Bound grammatical morphemes							
Δ Total words	.903	.816	.807	.816	92.956	1, 21	.000
No. of Spanish courses	.911	.829	.812	.013	1.556	1, 20	.227
Phonological memory	.911	.829	.802	.000	0.002	1, 19	.962
Δ Subordinate clause use							
Δ Total words	.725	.526	.503	.526	23.276	1, 21	.000
No. of Spanish courses	.823	.678	.646	.152	9.465	1, 20	.006
Phonological memory	.824	.678	.628	.000	0.022	1, 19	.885
Δ Coordinate clause use							
Δ Total words	.847	.717	.703	.717	53.127	1, 21	.000
No. of Spanish courses	.852	.725	.698	.009	0.635	1, 20	.435
Phonological memory	.853	.727	.684	.002	0.107	1, 19	.747
Δ Function words							
Δ Total words	.955	.911	.907	.911	215.241	1, 21	.000
No. of Spanish courses	.955	.912	.903	.000	0.100	1, 20	.755
Phonological memory	.956	.914	.901	.003	0.621	1, 19	.441
Δ Pronouns							
Δ Total words	.490	.240	.204	.240	6.621	1, 21	.018
No. of Spanish courses	.520	.270	.197	.030	0.827	1, 20	.374
Phonological memory	.525	.275	.161	.005	0.139	1, 19	.713
High Ability ($n = 20$)							
Δ Narrative abilities							
Δ Total words	.302	.091	.040	.091	1.802	1, 18	.196
No. of Spanish courses	.388	.150	.050	.059	1.185	1, 17	.291
Phonological memory	.421	.177	.023	.027	0.521	1, 16	.481
Δ Vocabulary							
Δ Total words	.857	.735	.721	.735	50.002	1, 18	.000
No. of Spanish courses	.879	.772	.745	.037	2.760	1, 17	.115
Phonological memory	.880	.774	.732	.002	0.123	1, 16	.730

Table 5 (*cont.*)

Variable	R	R^2	Adj. R^2	ΔR^2	ΔF	df	p
High Ability ($n = 20$)							
Δ Free grammatical morphemes							
Δ Total words	.558	.311	.273	.311	8.124	1, 18	.011
No. of Spanish courses	.566	.321	.241	.010	0.240	1, 17	.630
Phonological memory	.616	.379	.262	.058	1.503	1, 16	.238
Δ Bound grammatical morphemes							
Δ Total words	.557	.310	.272	.310	8.087	1, 18	.011
No. of Spanish courses	.560	.313	.233	.003	0.082	1, 17	.778
Phonological memory	.576	.332	.207	.019	0.449	1, 16	.512
Δ Subordinate clause use							
Δ Total words	.187	.035	−.018	.035	0.656	1, 18	.429
No. of Spanish courses	.393	.155	.055	.120	2.407	1, 17	.139
Phonological memory	.543	.295	.163	.140	3.185	1, 16	.093
Δ Coordinate clause use							
Δ Total words	.538	.289	.250	.289	7.319	1, 18	.014
No. of Spanish courses	.540	.291	.208	.002	0.051	1, 17	.824
Phonological memory	.540	.291	.158	.000	0.000	1, 16	.988
Δ Function words							
Δ Total words	.540	.291	.252	.291	7.405	1, 18	.014
No. of Spanish courses	.580	.337	.259	.045	1.157	1, 17	.297
Phonological memory	.703	.494	.399	.157	4.979	1, 16	.040
Δ Pronouns							
Δ Total words	.157	.025	−.030	.025	0.455	1, 18	.509
No. of Spanish courses	.173	.030	−.084	.005	0.092	1, 17	.765
Phonological memory	.296	.087	−.084	.057	1.007	1, 16	.331

by Gathercole et al. and French was vocabulary *knowledge*, whereas our study measured development of vocabulary *use*, suggesting that, perhaps, PM is not related to the development of productive vocabulary but, rather, to the development of vocabulary knowledge. Unlike young children who would use all or most of their vocabulary knowledge in speech, the productive vocabulary of adults in a given sample of speech most probably understates their vocabulary knowledge.

Second, the difference in results may be due to the PM measure used. Most studies, and certainly the studies cited above, have used nonword repetition as the measure of PM. Serial nonword recognition, the task used in the present study, is thought to primarily involve phonological storage (Baddeley, 2003), whereas nonword repetition involves at the very least both phonological storage and an articulatory component (Baddeley, 2003; see also Bowey, 2001, and Snowling, Chiat, & Hulme, 1991). Perhaps vocabulary acquisition is more closely related to the articulatory component of PM than to the nonarticulatory storage component measured by serial nonword recognition. The present data, however, do not permit us to differentiate between these two possibilities. Whether serial nonword recognition and nonword repetition make different predictions in vocabulary development should be investigated in future studies.

PM and narrative abilities

The present study found that PM was related to development of narrative abilities in the low ability participants such that those with better PM skills demonstrated greater narrative gains than those with weaker PM skills. Interestingly, PM did not predict development of narrative abilities in the high ability group, even though the groups did not differ on this measure either at Time 1 or Time 2. By definition (Biber, 1988), narrative abilities as measured here required the use of more complex grammar (past tense verbs, third-person morphology, past participles, present participles, and public verbs). The association between PM and narrative gains obtained for the low ability participants was consistent with developmental studies showing that preschool children with better PM skills produce more grammatically complex utterances (Adams & Gathercole, 1995, rs = .35 and .36; Adams & Gathercole, 2000, rs = .39–.51) and are better at repeating complex sentences (Blake et al., 1994; Willis & Gathercole, 2001) than children with poorer PM skills.

The verbs used in the narrative discourse constructions were also lexical items. Unlike the other grammatical elements investigated in the present study, the participants were credited with a narrative item *regardless of whether it was used correctly*. In his processing theory of L2 learning, VanPatten (2004) posits that early L2 learners process content words (lexical items) before they process noncontent words (nonreferential grammatical forms; principle 1a, p. 8). Thus, the narrative gains made by the low ability participants may have reflected essentially vocabulary development rather than or in addition to grammatical development. PM is known to relate to vocabulary acquisition (see Baddeley et al., 1998), and this finding is consistent with previous research looking at vocabulary development.

The failure to find a relationship between PM and narrative gains in participants demonstrating relatively high ability at the beginning of the study (those with higher SAT II Spanish scores) may be due to having a larger store of complex grammatical templates (e.g., high-frequency formulaic expressions such as *me dijo que*, *no sé si*), and hence, essentially lexical abilities that afford the generation of more complex utterances. Speidel (1989, 1993) proposed that children acquire syntactic complexity in their L1 by imitating adult models. The heard phrases are held initially in PM and subsequently transferred into long-term memory. Children eventually construct correct syntactic speech from the corpus of templates stored in long-term memory. Good PM skills are required to imitate adult utterances, especially if they are longer or more complex.

This model of grammar and syntax acquisition may well apply to adults learning an L2. Under this model, those with good PM skills would be able to hold longer and more complex utterances in PM, which would then be transferred to long-term memory. The stored representations would provide the basis for constructing grammatically correct speech. Those with poorer PM skills who would be able to hold only shorter and less complex utterances in PM would fail to build an adequate store of complex templates.

Good grammar skills reduce dependence on PM (see Romani, 1994; see also Blake et al., 1994). The high ability participants, presumably having greater knowledge of the L2 grammar, would place less reliance on PM. If grammatical ability is inefficient, as in young children learning their L1 or adults in the earlier phases of learning an L2, more reliance needs to be placed on PM. Consequently, for the

low ability participants who possessed less well-developed grammatical skill, PM played a more important role in narrative discourse.

PM and grammatical morphemes

This study showed that, after controlling for change in total number of words used, PM was related to the gain in the correct use of function words for the high but not for the low ability participants. This result was unexpected given the evidence that PM is implicated primarily at earlier stages of language learning. However, according to VanPatten's (2004) primacy of content words principle (principle 1a), and similar to children learning their L1 (Slobin, 1985), L2 learners process content words first and tend to pass over function words. This primary focus on meaning consumes working memory capacity. As proficiency increases, however, lexical retrieval becomes easier, freeing up working memory capacity to process nonmeaningful grammatical forms, such as function words and inflections. Hence, we should expect to see a relationship between PM and function words for the high ability participants, for whom lexical access would be easier.[2]

The finding that PM was associated with gains in the correct use of function words for the high but not the low ability participants is also consistent with VanPatten's (2004) sentence location principle (principle 1f). According to this principle, L2 learners process items in initial positions in a sentence, then the items in final positions, and last, those items in medial positions. Because function words usually appear in medial positions, such words would tend to be processed more by the more proficient (high ability) rather than the less proficient (low ability) learners. Of course, this relationship should be tested in a further study.

PM and clause use

PM was found to relate to the gain in subordinate clause use for high but not low ability participants. Once total words were partialed out, PM accounted for a further 14% of variance that showed a trend toward significance ($p = .093$), most probably due to a ceiling effect in the number of subordinate clauses produced by the high ability participants (see Table 1). Again, this result was unexpected. However, because of its inherent complexity, subordinate clause production is a grammatical skill that would emerge later in language (L1 and L2) development. For the more advanced learners (here the high ability participants), good PM skills would be required to create an adequate store of templates from which to construct utterances containing subordinate clauses. Use of subordinate clauses is highly related to the use of function words. It follows, then, that if PM is implicated in the correct use of function words for high ability participants, it should also be related to subordinate clause use for the same participants.

PM did not account for a significant amount of variance in coordinate clause use for either group. The reason that there was no relationship between PM and coordinate clause use may be because increased use of coordinate clauses does not mean increased sentence complexity. In effect, coordinate clauses are simply two sentences connected by a conjunction. The present results are consistent with the failure of Willis and Gathercole (2001) to find PM group differences for the repetition of easy sentences.

The role of vocabulary in the relationship between PM and grammatical production

Some research has suggested that gains in L2 skills are mediated by vocabulary *knowledge* (French, 2004a; Service & Kohonen, 1995). Our data, reflecting the number of unique *vocabulary items produced* rather than vocabulary knowledge, did not permit us to address this question directly. However, we did have a measure for the total number of words participants produced in their speech samples. There was a clear dissociation between the low and high ability participants in terms of the amount of variance that was explained by total words. For the low ability participants, the gain in total words explained between 24 and 91% of the variance in grammatical gains, whereas for the high ability participants, the variance explained by total words was only between 3 and 31% (*mostly, nonsignificant*) for grammatical measures. For the low but not the high ability participants, grammatical production was highly driven by the amount of speech they produced. For the high ability participants, grammatical production appeared to be less reliant on the amount of speech produced. That is, they were able to increase their use of grammatical elements without a corresponding increase in speech output. The results of the present study suggest that L2 skills are mediated by the volume of speech output primarily at early stages of learning.

PM as a mediator of grammar acquisition

In addition to learning individual words, PM appears to be important in learning the sentence order of words. Both Romani (1994) and Speidel (1993) describe individuals with impaired PM skills whose syntactic deficits primarily involved misplacement or omission of function words and word order problems. The correct use of function words requires not only learning the word itself, but that it be placed in the correct order in the sentence. Ellis and Sinclair (1996) propose that learning of syntax is similar to learning words. When learning words, the phonemes must be placed in the correct order. In learning syntax, the words must be placed in correct sequence. Such sequences, be they phonemes or words, are initially held in PM. They suggest that grammar acquisition is simply the generalization of such sequences.

CONCLUSION

This study demonstrated that PM is implicated in the development of narrative abilities for learners whose initial ability with the language is relatively low (as evidenced by the 17.5% of unique variance explained by PM for gain in narrative abilities for the low ability group) and in the development of accurate use of function words in learners with higher initial ability (as evidenced by the 15.7% of unique variance explained by PM for gain in the correct use of function words for the high ability group). This study also showed that at lower levels of achievement, narrative abilities and grammatical competence is heavily driven by the volume of speech output. At higher levels of ability, learners are able to use more complex grammatical elements without a corresponding increase in speech output. Although some previous research has suggested that PM is important in the language development of younger learners, the present study extends this finding

by demonstrating that PM is implicated in the language development of adult L2 learners, at stages when language learning becomes more demanding (effortful).

It is suggested that at early stages of language learning, when grammatical knowledge is weak, the learner concentrates on learning content words. For these learners, grammar acquisition is heavily driven by vocabulary acquisition for which PM is important. At this stage, vocabulary learning consumes all or most of the learner's PM capacity. At later stages of language learning when lexical access is easier, PM capacity is redeployed for learning more complex grammar. At even later stages, fluent speakers, whether they are L1 or L2 speakers, should minimize their dependence on PM, as language production would have become virtually automatic.

In conclusion, the results presented herein demonstrate that PM is implicated in the development of oral L2 skill in adults. The present study represents an initial investigation into the relationship between PM and L2 speech production in adults, and the failure to find a relationship between PM and some of the measures may have been due to a lack of power. Thus, these results need to be extended in future investigations using a larger number of participants.

APPENDIX A

The nonwords used in the serial recognition task

Four-item lists (practice)	bɑɹtʃ tig nʌp gʊk
	kɪb dɚn pʌtʃ gid
	mɑt tʃɛn pɛd kɪg
	mɚd gɑɹp tæm pɪb
Five-item lists	tɛk kæm mɪtʃ bɑn dɚp
	tʃum kɚp lɔɹk nʌg gɑɹn
	pim tɑɹg gæb bɑk tʃɛl
	gʊt bɑɹg mɚn nʌk tɛp
	lʊg dʒɑl ditʃ kɚm mɛb
	kitʃ dʒɑɹn mɛp tɚg bɪk
	kum tɔɹd mɑdʒ dʒʌp gɪk
	lɑd tʌdʒ dʒɪk nɔɹb gɑɹm
Six-item lists	bɔɹdʒ tʃʌd nig dæk kɛb lɑɹm
	tɚdʒ dʒup lɛk nɔɹg tʃɪm pib
	pɑɹn mæb dɔɹdʒ nɑg tʃim dʒɪt
	kɚn budʒ tʌd lɪg pæb dɔɹt
	tɪdʒ mʌp tʃɛn gɚb nʊg dit
	tɔɹm pæg jɛk dɚb cɑl bʌp
	kɔɹp tib nul dʒɑɹk pɪm gɚtʃ
	pædʒ nɑɹp mɑn tʃʌt gub gɛd
Seven-item lists	tuk tʃæd lɪdʒ dʒɚg dɑp nɑɹt gʌb
	kɑɹk nɚtʃ mɔɹd tʃæm bul lʌb tip
	tʃɑɹn nɪg kɑm dʒil gædʒ lɚb mʌn
	tɑb gæn dɑɹtʃ tʃul dʒʌk nɔɹd pɛm
	gɛl nɚg lʌd pɑk mitʃ dub dʒæt
	lim kʌg tʃɔɹdʒ dʒɚt kɛd dɑ́ɹp gɑk
	tʃɪg næm pɛb gɑp dʒutʃ lɑɹt tid
	dʒɑɹn nɛb gɚp tʃɔɹg mæl tutʃ lɑn

ACKNOWLEDGMENTS

This study was conducted by the first author as part of her doctoral dissertation research at the Université du Québec à Montréal, supervised by the second author (N.S.). The research was funded by a grant to the fourth author (B.F.) from the Council for International Educational Exchange (New York); by a grant from the Natural Sciences and Engineering Research Council of Canada to the second author (N.S.); by support from the Dean's Office, Faculty of Arts and Science at Concordia University, to the second author (N.S.); and in part by the Fonds de Recherche en Santé du Québec to the first author (I.O.). The authors thank Manuel Díaz-Campos, Barbara Lafford, and Nicole Lazar, who were members of a research team involved in a larger project of which this was one part. The authors also thank Peter Scherzer and Leif French for their helpful comments on this paper. Finally, the authors are grateful to Conchita Bueno, Hazel Casas, Elizabeth Gatbonton, Randall Halter, Guy Lacroix, Anne-Marie Linnen, Magnolia Negrete, Laura Renteria-Díaz, Marlene Taube-Schiff, Pavel Trofimovich, and Naomi Yamasaki, who helped during various phases of this project.

NOTES

1. The OPI is a standardized means for evaluating overall speaking ability. In this study, it was used as a method for collecting speech corpora.
2. This study was part of a larger project looking into differences between Study Abroad and At Home regular classroom (in the students' home country) learning contexts (Segalowitz & Freed, 2004; see also Collentine, 2004; Díaz-Campos, 2004; Lafford, 2004; Lazar, 2004). Interestingly, among the tasks reported in these papers was a Spanish lexical access task using well-known concrete nouns (see Segalowitz & Freed, 2004, for details). The high ability group (Time 1: $M = 14.30$, $SE = 1.17$; Time 2: $M = 10.95$, $SE = 0.79$) made significantly fewer errors on this task than the low ability group, Time 1: $M = 17.91$, $SE = 0.87$; Time 2: $M = 17.17$, $SE = 1.02$; $F(1, 41) = 18.10$, $MSE = 28.58$, $p < .001$, confirming that lexical retrieval was easier for this group than for the low ability group.

REFERENCES

Adams, A.-M., & Gathercole, S. E. (1995). Phonological working memory and speech production in preschool children. *Journal of Speech and Hearing Research, 38*, 403–414.

Adams, A.-M., & Gathercole, S. E. (1996). Phonological working memory and spoken language development in young children. *The Quarterly Journal of Experimental Psychology, 1996*, 216–233.

Adams, A.-M., & Gathercole, S. E. (2000). Limitations in working memory: Implications for language development. *International Journal of Language and Communication Disorders, 35*, 95–116.

Atkins, P. W. B., & Baddeley, A. D. (1998). Working memory and distributed vocabulary learning. *Applied Psycholinguistics, 19,* 537–552.

Baddeley, A. (1986). *Working memory.* Oxford: Clarendon Press.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*, 417–423.

Baddeley, A. D. (2003). Working memory and language: An overview. *Journal of Communication Disorders, 36*, 189–208.

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

Baddeley, A., & Hitch, G. (1974). Working memory. In G. A. Bower (Vol. Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–89). New York: Academic Press.

Biber, D. (1988). *Variation across speech and writing*. New York: Cambridge University Press.

Blake, J., Austin, W., Cannon, M., Lisus, A., & Vaughan, A. (1994). The relationship between memory span and measures of imitative and spontaneous language complexity in preschool children. *International Journal of Behavioral Development, 17*, 91–107.

Bowey, J. A. (2001). Nonword repetition and young children's receptive vocabulary: A longitudinal study. *Applied Psycholinguistics, 22,* 441–469.

Breiner-Sanders, K., Lowe, P., Miles, J., & Swender, E. (2000). ACTFL proficiency guidelines: Speaking, revised 1999. *Foreign Language Annals, 33*, 13–18.

Cheung, H. (1996). Nonword span as a unique predictor of second-language vocabulary learning. *Developmental Psychology, 32,* 867–873.

Collentine, J. (2004). The effects of learning contexts on morphosyntactic and lexical development. *Studies in Second Language Acquisition, 26*, 227–248.

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.

Díaz-Campos, M. (2004). Context of learning in the acquisition of Spanish second language phonology. *Studies in Second Language Acquisition, 26*, 249–273.

Ellis, N. C., & Sinclair, S. G. (1996). Working memory in the acquisition of vocabulary and syntax: Putting language in good order. *The Quarterly Journal of Experimental Psychology, 49A*, 234–250.

French, L. (2004a). Phonological working memory and L2 acquisition: A developmental study of Québec Francophone children learning English (Doctoral dissertation, Université Laval, 2003). *Dissertation Abstracts International, 65(2-A)*, 487.

French, L. (2004b, May). *Does phonological memory ability predict gains in children's L2 grammatical knowledge?* Paper presented at the Annual Conference of the American Association for Applied Linguistics, Portland, OR.

Gathercole, S. E. (1998). The development of memory. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 39*, 3–27.

Gathercole, S. E., & Pickering, S. J. (1999). Estimating the capacity of the phonological short-term store. *International Journal of Psychology, 34*, 378–382.

Gathercole, S. E., Pickering, S. J., Hall, M., & Peaker, S. M. (2001). Dissociable lexical and phonological influences on serial recognition and serial recall. *The Quarterly Journal of Experimental Psychology, 54A,* 1–30.

Gathercole, S. E., Willis, C. S., Ellis, H., & Baddeley, A. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental Psychology, 28*, 887–898.

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

Hulme, C., Maughan, S., & Brown, G. D. A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language, 30,* 685–701.

Hulme, C., Thomson, N., Muir, C., & Lawrence, A. (1984). Speech rate and the development of short-term memory span. *Journal of Experimental Child Psychology, 38*, 241–253.

Karakas, S., Yalin, A., Irak, M., & Erzengin, Ö. U. (2002). Digit span changes from puberty to old age under different levels of education. *Developmental Neuropsychology, 22*, 423–453.

Lafford, B. A. (2004). The effect of the context of learning on the use of communication strategies by learners of Spanish as a second language. *Studies in Second Language Acquisition, 26*, 201–225.

Lafford, B. A., & Collentine, J. G. (1987). Lexical and grammatical access errors in the speech of intermediate/advanced level students of Spanish. *Lenguas Modernas, 14*, 87–112.

Lafford, B. A., & Collentine, J. G. (1989). The telltale targets: An analysis of access errors in the speech of intermediate students of Spanish. *Lenguas Modernas, 16*, 143–162.

Lazar, N. A. (2004). A short survey on causal inference, with implications for context of learning studies of second language acquisition. *Studies in Second Language Acquisition, 26*, 329–347.

Masoura, E. V., & Gathercole, S. E. (1999). Phonological short-term memory and foreign language learning. *International Journal of Psychology, 34*, 383–388.

Masoura, E. V., & Gathercole, S. E. (2005). Contrasting contributions of phonological short-term memory and long-term knowledge to vocabulary learning in a foreign language. *Memory, 13*, 422–429.

Nimmo, L. M., & Roodenrys, S. (2002). Syllable frequency effects on phonological short-term memory tasks. *Applied Psycholinguistics, 23,* 643–659.

O'Brien, I., Segalowitz, N., Freed, B., & Collentine, J. (2006). *Phonological memory predicts second language oral fluency in adults.* Manuscript submitted for publication.

Romani, C. (1994). The role of phonological short-term memory in syntactic parsing: A case study. *Language and Cognitive Processes, 9*, 29–67.

Segalowitz, N., & Freed, B. F. (2004). Context, contact and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisition, 26*, 173–199.

Service, E. (1992). Phonology, working memory, and foreign-language learning. *The Quarterly Journal of Experimental Psychology, 45A*, 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.

Slobin, D. I. (1985). Crosslinguistic evidence for the language-making capacity. In D. I. Slobin (Seri. & Vol. Ed.), *The crosslinguistic study of language acquisition* (pp. 1157–1256). Hillsdale, NJ: Erlbaum.

Snowling, M., Chiat, S., & Hulme, S. (1991). Words, nonwords, and phonological processes: Some comments on Gathercole, Willis, Emslie, & Baddeley. *Applied Psycholinguistics, 12*, 369–373.

Speciale, G., Ellis, N. C., & Bywater, T. (2004). Phonological sequence learning and short-term store capacity determine second language vocabulary acquisition. *Applied Psycholinguistics, 25*, 293–321.

Speidel, G. E. (1989). Imitation: A bootstrap for learning to speak? In G. E. Speidel & K. E. Nelson (Eds.), *The many faces of imitation in language learning* (pp. 151–179). New York: Springer–Verlag.

Speidel, G. E. (1993). Phonological short-term memory and individual differences in learning to speak: A bilingual case study. *First Language, 13*, 69–91.

The College Board. (2000). *Taking the SAT II: Subject test*. New York: Author.

Thorn, A. S. C., Gathercole, S. E., & Frankish, C. R. (2002). Language familiarity effects in short-term memory: The role of output delay and long-term knowledge. *The Quarterly Journal of Experimental Psychology, 55A,* 1363–1383.

VanPatten, B. (2004). Input processing in second language acquisition. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 5–31). Mahwah, NJ: Erlbaum.

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

Willis, C. S., & Gathercole, S. E. (2001). Phonological short-term memory contributions to sentence processing in young children. *Memory, 9*, 349–363.