

The role of reading span in factual and inferential comprehension and retention in L2 reading*

Hyangsook Park · Kichun Nam · Yae-Sheik Lee**

(Kyungpook National University · Korea University · Kyungpook National University)

Park, Hyangsook, Kichun Nam, and Yae-Sheik Lee. 2016. The role of reading span in factual and inferential comprehension and retention in L2 reading. *Linguistic Research* 33(Special Edition), 81-106. The present study attempts to explain the relationship of working memory (WM) to L2 reading comprehension (RC) by conducting two experiments. The participants' WM was measured by two reading span (RS) tests adapted from Daneman and Carpenter (1980). Specifically, Experiment 1 explores the role of L1-RS and L2-RS in overall RC in L2. Fifty-eight Korean undergraduate students participated in this experiment. The results showed that L2-RS had better predictive power than L1-RS for L2 RC performance and that the participants performed better in the L1-RS than L2-RS test. Experiment 2 further investigates the role of RS in terms of types of comprehension and retention. For this, fifty Korean students were divided into two groups according to their L2-RS. The RC tests were given in two types of questions: factual and inferential. The retention of information was assessed by two memory tests: immediate and delayed tests. The results revealed that the high-RS group consistently outperformed the low-RS group on the two types of RC questions in the two memory tests. No significant decay, however, was found between the memory tests. In addition, there was no interaction between group, question type, and time. The findings showed that RS does play a significant role in L2 RC performance. (Kyungpook National University · Korea University)

Keywords working memory, reading span, L2 reading comprehension, mental representation, factual and inferential comprehension, retention, language learning

1. Introduction

Reading comprehension (RC) involves complex cognitive processes, including

* This article is based on Hyangsook Park's doctoral dissertation (2015) completed at Kyungpook National University. We would like to thank anonymous reviewers for their valuable comments and suggestions.

** Corresponding author

decoding, understanding, and maintaining information in memory. The capacity for processing and storage of information is limited, so working memory (WM) plays an important role in RC. Working memory (WM) is a dynamic mechanism that stores information temporarily while it is also engaging in other cognitively demanding tasks (Baddeley 2000, 2003). Among WM measures, reading span (RS) is a good predictor for RC performance. For instance, RS reflects the degree to which a person can maintain information in short-term memory while another processing task is executed simultaneously. RS reflects the capacity of the phonological loop and central executive of WM (e.g., Daneman and Carpenter 1980; Daneman and Merikle 1996).

A substantial body of research has attempted to explain inter-learner variation in L1 reading comprehension performance by RS (e.g., Burton and Daneman 2007; Daneman and Merikle 1996). Few studies, however, have examined the role of RS in L2 RC, and the findings are mixed (e.g., Gass and Lee 2011; Juffs 2004; Osaka and Osaka 1992). Therefore, it lacks the evidence that can claim that the same is true for L2 reading comprehension.

Thus, the present study attempts to obtain a more comprehensive understanding of the RS role in L2 RC performance by conducting two experiments. Specifically, Experiment 1 explores the association between L1-RS and L2-RS and the role of RS in overall L2 RC performance. Experiment 2 further examines the association between RSs and L2 RC by taking into account both factual and inferential comprehension on the basis of Kintsch's (1990) mental representations (surface, textbase, and situational model).

2. Literature review

2.1 Reading span

Working memory (WM) is “the temporary storage and manipulation of information that is assumed to be necessary for a wide range of complex cognitive activities” (Baddeley 2003: 189). Baddeley (2000) introduces four components of WM system: the central, phonological loop, visuospatial sketchpad, and episodic buffer.¹ According to the theory of WM, there are a limited number of things people can do at once, and the amount of information that can be processed and stored

simultaneously varies among individuals (Carroll 2008).

Various WM tasks are generated to measure WM capacity, such as simple or complex WM tests.² Among WM tests, RS is considered to be a factor that can explain inter-learner variation in reading comprehension performance. RS tests force the participants to operate both functions of storage and processes of WM (Daneman and Carpenter 1980). There is a trade-off between the two functions due to the limitation of WM. For example, a task demanding more processing time may leave less capacity for storing information. The present study uses an RS test to measure WM capacity, so that these two terms will be interchangeably used later on.

2.1.1 Relationship between L1 and L2 reading spans

A body of research supports that there is a significant correlation between L1-RS and L2-RS, but L2-RS is influenced by the students' L2 proficiency (e.g., Alptekin and Erçetin 2010; Coughlin and Tremblay 2013; Harrington and Sawyer 1992; Osaka *et al.* 1993; Service *et al.* 2002). Unambiguously, L2 processing is more demanding compared with L1 processing. Students' speed of articulation may slow down in the L2-RS test, which reflects less efficient functions of the phonological loop. Drops in articulation speed lead them to more consumption of resources, which in turn may cause difficulty rehearsing and remembering target words. With an increase in L2 proficiency, the gap between the two RSs becomes smaller (Ortega 2009). For example, Osaka *et al.* (1993) found that despite the significant correlation between L1-RS and L2-RS, the participants obtained higher scores on RS test in the L1 (German) than in the L2 (French). Similarly, Service *et al.*'s study (2002) showed that those with less proficiency in their L2 received higher RS in their L1 (Finnish) than L2 (English). However, there were no differences for the more advanced participants

¹ The phonological loop, a model of short-term memory (STM), stores verbal information through rehearsal. The visuospatial sketchpad handles visual and spatial information. The central executive actively processes and controls the two slave systems, phonological loop and visuospatial sketchpad, and allocates attentional resources to the task. The episodic buffer combines information from the loop, the sketchpad, and long-term memory (LTM), into a coherent episode. It also provides a link between the central executive and LTM. Think of it this way. The central executive has a limited capacity and delegates the processing syntactic and semantic information as a way of managing the flow of information within the tight confines of the processing system (Baddeley 2000, 2003).

² Simple WM tests measure passive storage capacity and are based on the number or span of unrelated digits or words recalled. Complex WM tests (RS) assess both storage and processing.

in their L2. Van den Noort *et al.* (2006) also suggest that underlying WM capacity in the L1 (Dutch) and L2 (German) are closely related and that the participants' WM capacity increased in accordance with their proficiency level in the L2.

As noted above, the majority of studies used languages which share similar semantic and syntactic features: English and Spanish, English and French, Dutch and German. Relatively few studies have compared L1-RS and L2-RS which share few linguistic similarities: English and Japanese (e.g., Osaka and Osaka 1992). For instance, Osaka and Osaka's (1992) within-subject study compared Japanese (L1-RS) and English (L2-RS). They administered the RS tests adapted from Daneman and Carpenter's (1980). For the L2-RS test, the participants were asked to recall the last word of each sentence within each set. For the L1-RS test, the participants were asked to recall the word underlined in red, since the position of the target word was randomized using different parts of speech. The study found high correlations between the L1-RS and L2-RS, suggesting that RS tests in both languages are valid to predict reading efficiency. In addition, the participants performed better on their L1-RS than their L2-RS tests.

2.2 Reading span and reading comprehension

2.2.1 Reading span and overall reading comprehension

Successful comprehension of texts requires various skills and knowledge, such as word-decoding, syntactic skills, and vocabulary knowledge (Daneman and Carpenter 1980; Waters and Caplan 1996). For example, the reader should simultaneously process the information from what is being read and integrate it with other incoming information. A body of research on L1 RC has shown that RS is a reflection of limitations of cognitive capacity (e.g., Daneman and Carpenter 1980). For instance, Daneman and Carpenter's (1980) landmark study on the role of RS in L1 RC found that skilled readers had a greater RS. That is, large capacity facilitates efficient processing which gives the reader more available resources to integrate facts into the general representation. In addition, large capacity enhances consolidation³ of the newly obtained information. This memory consolidation increases the availability in

³ Consolidation is a process that converts information from STM into LTM (Dudai 2004).

subsequent processing and later accessibility of the information.

In contrast, few studies have investigated the relationship of RS (L1, L2, or both) in L2 sentence-level comprehension and yielded mixed results: correlations (e.g., Coughlin and Tremblay 2013; Miyake and Friedman 1998; Sagarra 2007; Sagarra and Herschensohn 2010) and no correlations (e.g., Felser and Roberts 2007). For example, Sagarra (2007) recruited English learners of Spanish of a low proficiency level and examined the relationship between L1-RS and sensitivity to gender agreement in Spanish. The results revealed that the students with a high-RS received higher scores than those with a low-RS on the gender agreement test. Likewise, Sagarra and Herschensohn (2010) recruited English learners of Spanish at a different proficiency level, an intermediate level, and yielded similar results in which the high-RS learners received higher scores than the low-RS readers on the test. Coughlin and Tremblay (2013) examined the effects of L2-RS in processing L2 French number agreement by recruiting L1 English speakers. The findings also showed that such performance was affected by RS. Harrington and Sawyer (1992) included both L1-RS (Japanese) and L2-RS (English) and explored whether RSs are related to individual differences in L2 RC performance. The results showed that those with a high-RS performed better than those with a low-RS on the test.

In contrast, Juffs' (2004) study recruited students from three different L1 backgrounds (Chinese, Japanese, and Spanish) and investigated the role of RS in L2 (English) processing of temporary syntactic ambiguity. Juffs found no significant role of RS in processing L2 syntactic sentences. Similarly, Felser and Roberts' (2007) study also found no significant differences in comprehending English *wh*-sentences between high and low-RS Greek learners.

2.2.2 Reading span and mental representations

The reader actively interacts with the text and builds mental images of the content while reading, which is called mental representation. Kintsch (1990) argues that knowing the content explicitly stated in the text is not sufficient for successful comprehension of a text. The reader should integrate information obtained from the text into some larger structure by using prior knowledge and constructing correct situational models. Reading involves a continuous process of information and the use of metacognitive strategies when constructing meaning from texts. Automaticity

reduces WM load on comprehending a text and allows readers to readily retrieve information stored in their LTM. Achieving automaticity and fluency, therefore, is vital to become efficient readers (e.g., Service *et al.* 2002).

Several studies have shown that RS plays a critical role in RC (e.g., Estevez and Calvo 2000). For instance, Kintsch *et al.* (1990) investigated how three levels of comprehension⁴ in L1 texts are retained in memory. The participants were placed into one of four test conditions (right after reading, after delays of 40 min, 2 days, and 4 days) and instructed to complete a recognition test based on the three-level comprehension. The results revealed that there were significant differences in the decay rates among the three levels. For instance, surface-level memory was observed only on the immediate test and rapidly decreased with delay. Memory for textbase-level was strong on the immediate test, but it also decreased over time. However, there was no significant decline in memory at the situational model-level. Estevez and Calvo (2000) found that low-RS readers had short response time with lower accuracy rates than high-RS readers. The findings suggest that high-RS may enhance consolidation of memory in that the reader can have more available resources for deep processing at the situational model-level.

Relatively few studies on RS have examined its relationship to L2 comprehension in terms of the levels of representations. L2 learners do not have the same level of linguistic competence as that of their L1. Inevitably, they are likely to consume more resources in the surface-level comprehension, which in turn means they have less available resources for deep processing (Perkins and Brutton 1992; Walter 2004).

In a similar vein, Alptekin and Erçetin (2009) propose that the role of RS in RC is more evident when tasks are more challenging. For instance, literal understanding relies on surface features, such as decoding and syntactic parsing. Generating inferences, in contrast, requires more cognitive resources than literal understanding. They recruited thirty Turkish learners of English from a university in Turkey and examined the role of RS (recognition and recall tests) in L2 literal and inferential comprehension. Not surprising, they found a significant relationship only between

⁴ *Surface*: verbatim memory of the words, phrases, and sentences of actual texts.

Textbase: a coherent conceptual representation of a text and its structure (i.e., propositions, summary of the text).

Situational Model: the model of the situation described by the text: integration of the new text and prior knowledge or creating links between them (i.e., inference, elaboration).

inferential comprehension and recall-based RS. Alptekin and Erçetin (2010) confirmed their previous findings by yielding a significant correlation between RS and generating inferences. The results of the two studies indicate that those with a high-RS can process information more efficiently and thus have more available resources for generating inferences and deep-level understanding (e.g., Engle and Marshall 1983). That is, the more cognitive resources they can access, the better able they are to process input and learn (e.g., Alptekin and Erçetin 2010; McLaughlin 1995; Wen 2014; Waters and Caplan 1996).

2.3 The difference between Korean and English

There are several difficulties for Korean learners of English which stem from a number of differences between the two languages (Baek 2013; Kim, 2015; Pae 2011). First of all, Korean learners should familiarize themselves with English stress and rhythm since Korean is a syllable-timed language in which each syllable is given the same emphasis in a word. Second, there is a difference in the word order between the two languages. For example, Korean is a head-final (subject-object-verb) language that Korean sentences end with verb phrases. In contrast, English sentences end with noun phrases since English is a head-initial (subject-verb-object) language. Lastly, Korean is an agglutinative language in which particles and suffixes are added one after the other to a stem. For example, a noun in a sentence is attached one or more particles to the end of it (e.g., Kangaci-*ka*: puppy SUBJ). In addition, verb information, such as tense and mood, is added to the end of the verb: nulk-*nunta*: to age suffix (Grady 1991). Taken together, Korean learners of English are likely to encounter difficulties in comprehending English in that a great degree of cognitive restructuring is required. Thus, it can be assumed that RS plays a critical role in successful L2 use (e.g., Dijkstra 2005).

There are several issues to be considered when investigating the role of RS in L2 RC performance. First, the reader's proficiency levels are not controlled. Second, the findings are mixed and lack evidence to support the findings due to the dearth of studies (e.g., Juffs 2004). Third, relatively few studies have included both L1-RS and L2-RS, and where both languages (i.e., English and Spanish) have semantic and syntactic similarities. Accordingly, it is unknown whether the degree of difference between L1 and L2 can affect the association between L2 RC and RS. Finally, the

majority of RS studies on L2 RC have focused on word or sentence-level comprehension. Therefore, the present study attempts to explain individual differences in L2 RC among Korean learners of English whose L1 shares few linguistic similarities with L2 (English). To my knowledge, no known previous studies recruited Korean learners of English at a similar proficiency level. The participants' L2 RC performance was assessed in terms of both overall comprehension and two different types of comprehension (factual and inferential).

3. Two experiments

Two experiments were conducted to investigate the relationship of RS to L2 RC concerning overall and different levels of comprehension.

3.1 Experiment 1

Experiment 1 explores the relationship of L1-RS and L2-RS to overall L2 RC among Korean learners of English. A specific research question for the first experiment is as follows:

Are there any differences in RS scores between L1-RS and L2-RS tests? If so, how well do L1-RS and L2-RS predict L2 RC performance?

3.1.1 Methodology

3.1.1.1 Participants

Fifty-eight Korean university students (*Male* = 31, *Female* = 27) whose ages ranged from 20 to 28 (*Mean age* = 23.20 years) participated in the study. The participants had studied English as an L2 for at least 10 years through formal English education. The participants' proficiency levels were similar in terms of their TOEIC scores, ranging from 750 to 900 ($M = 833.93$). Therefore, the participants were homogeneous in terms of their L2 proficiency and educational background. The participants were recruited through university Internet bulletin boards and were paid for their participation.

3.1.1.2 Materials

3.1.1.2.1 Language background information questionnaire

A language background questionnaire was designed to collect the participants' personal information, including age, sex, age of first exposure to English, years of English language study, English study materials, and previous visits to English speaking countries.

3.1.1.2.2 Reading comprehension test

L2 RC ability of the participants was evaluated by a computerized RC test from a TOEFL practice test (2007). The test was composed of thirty-eight multiple-choice questions based on ten short texts (mean words per text = 111.1). Each text contained three or four questions, comprising main ideas, important details, and inferences.

3.1.1.2.3 Reading span tests

To measure the participants' RS, two versions of the RS test, Korean (L1) and English (L2), were constructed by adapting the design of Daneman and Carpenter (1980). Each version of the RS test consisted of five different levels, and each level contained five sets (from level two to six, a total of 100 unrelated sentences). Twenty sentences were used for a recognition test in order to prevent the participants from adopting a strategy of focusing on the target words without allocating much attention to reading the sentences (Daneman and Carpenter 1980). To meet the methodological criteria, the length of the sentences both in the L1 and L2 was controlled, ranging from eight to ten words for each sentence. The target words for the two RS tests were nouns, and the number of syllables in the target words was from two to three. The words used in the English RS test belonged to the first 5,000 word families of English based on frequency of occurrence in the British National Corpus of English. Semantic relations of the target words within each set were also controlled by avoiding using words with similar meaning. The participants were instructed to recall the first words (stems without suffixes) of each sentence within

the set for the L1-RS test, and the last word for the L2-RS test.⁵ The sentences in the L2-RS test were simple and active to avoid possible floor effects⁶ in performance because of task difficulty. The levels of difficulty of the sentences were evaluated by students at a similar proficiency level before the experiment, and they found that the difficulty of the sentences was almost the same.

3.1.1.2.4 Self-evaluation of reading span tests

A 5-point Likert scale self-evaluation questionnaire was designed to assess difficulties the participants experienced while they were taking the RS tests.

3.1.1.3 Procedure

The data were collected in a quiet classroom by testing the participants individually. The average experiment time for each participant took about 90 minutes. After the general instructions of the tests, the participants were instructed to complete the language background information questionnaire. Then they were asked to complete the computer-based RC test at their own pace. The participants could neither go back to the text section nor to a previous question to change their answer. The students were instructed to answer the comprehension questions with the text unavailable, which brings a significant reliance on the reader's RS (Alptekin and Erçetin 2009). The computer recorded the participants' answer choices and reading times. After completing the RC test, they were given a five-minute break.

After the break, each RS test started with five practice trials to familiarize the participants with the tests. The participants were instructed to read aloud each sentence at their own pace. It is assumed that reading out loud keeps them from covert rehearsal (repetitive practicing in their minds). Each sentence was presented for seven seconds. A cross appeared on the screen when each set had ended. Then they were instructed to recall the target words of each sentence in the set in the order in which they were presented. They completed the recognition test before advancing to the next level. The researcher recorded their responses on the answer sheets. In order to avoid

⁵ The rationale behind this is that unlike English, Korean sentences end with a verb.

⁶ A floor effect occurs when most of the subjects score near the bottom. There is very little variance because the floor of a test is too high.

order effects, half of the participants completed the L1-RS test first, and the other half completed the L2-RS test first. After the first RS test, the participants were given a five-minute break in order to reduce the effects of fatigue in the second RS test. After finishing RS tests, the participants were asked to rate the difficulty level by using a 5-Likert scale, 1 for the easiest and 5 for the most difficult.

3.1.1.4 Scoring

In scoring, the participants received one point for each correct answer on the RC test. For the RS tests, when the participants correctly recalled target words more than three target words of the total set of five at a particular level, the level was taken as their RS. If the participants recalled two target words correctly out of the five sets at the next level, the participants were given a half point. For instance, if a participant recalled three sets of target words correctly at the three-sentence level, and then recalled two sets of target words correctly at the four-sentence level, the participant's RS would be 3.5. For the self-evaluation of the RS tests, the score corresponded to the participants' number choice. For example, if the participants chose 5, the test was very difficult.

3.1.1.5 Data analysis

Stepwise regression analyses were conducted to evaluate predicted scores for L2 RC scores. The dependent variable was the RC scores, and the common predictor variables were L2 proficiency (TOEIC scores) and age of first exposure to L2. Two different predictor variables for each analysis were L1-RS and L2-RS.

3.1.2 Results and discussion

3.1.2.1 Descriptive statistics

The RC scores of only fifty-six participants were included in the data analysis because two participants did not have RC scores due to answer input error. The descriptive statistics for all measures used in the study are shown in Table 1. Table 2 displays the intercorrelation between the variables. The results revealed that the

participants obtained higher RS scores in the L1 than in the L2: $M = 3.98$ and 3.13 , respectively. The correlation between L1-RS and L2- RS was significant ($r = .524$, $p < .001$). The assumptions of linearity, normally distributed errors, and uncorrelated errors among the variables were checked and met.

Table 1. Descriptive statistics ($N = 56$)

	<i>M</i>	<i>SD</i>	<i>Minimum</i>	<i>Maximum</i>
<i>TOEIC Scores</i>	834.20	46.04	750	900
<i>RC Scores</i>	28.75	3.22	21	34
<i>L1-RS</i>	3.98	.69	2.5	6
<i>L2-RS</i>	3.13	.50	2.0	4.5
<i>Age of First Exposure to L2</i>	9.61	2.08	4	13

* RC: L2 reading comprehension, RS: reading span

Table 2. Intercorrelations: RC performance and predictor variables

	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>
<i>RC Scores</i>	.385**	.365**	.175	.426**
<i>Predictor</i>				
1. <i>L1-RS</i>	-	.524**	.091	.310*
2. <i>L2-RS</i>		-	.005	.110
3. <i>Age of 1st Exposure to L2</i>			-	.175
4. <i>L2 Proficiency</i>				-

*RS: reading span, RC: L2 reading comprehension * $p < .05$, ** $p < .001$

3.1.2.2 Predictive power of L1-RS and L2-RS in L2 reading comprehension

The results revealed that L2 proficiency, L1-RS, and L2-RS were important factors that influence L2 RC performance. In addition, L2-RS had a stronger predictive power for L2 RC performance than L1-RS, as shown in Table 3. Specifically, Model 2 explained 22.4% of the variance in L2 RC performance by eliminating the age of first exposure to L2 from Model 1: Adjusted $R^2 = .224$, $F(2, 53) = 8.95$, $p < .001$, with a medium effect (Cohen 1988). The beta weights show that proficiency ($\beta = .34$, $t = 2.72$, $p = .009$) and L1-RS ($\beta = .28$, $t = 2.24$, $p = .029$)

contributed to predicting L2 RC performance positively. Model 4 also explained 24.3% of the variance in L2 RC performance by eliminating the age of first exposure to L2: Adjusted R^2 of .241: $F(3, 52) = 9.82, p < .001$, with a medium effect. The beta weights show that proficiency ($\beta = .38, t = 3.15, p = .003$) and L2RS ($\beta = .30, t = 2.54, p = .013$) contributed to predicting L2 RC performance positively. VIF (Variation Inflation Factor) was not found in any multivariate analysis.

Table 3. Model summary: Predicting reading comprehension performance

<i>Model</i>	<i>Predictor</i>	<i>R</i>	<i>Adjusted R²</i>	<i>B</i>	<i>SEB</i>	<i>β</i>	<i>t</i>	<i>p</i>	<i>VIF</i>
1	<i>1st Exposure</i>	.51	.218	-.15	.196	-.09	-.77	.446	1.03
	<i>L2 Proficiency</i>			.02	.009	.32	2.55*	.014	1.13
	<i>L1-RS</i>			1.28	.581	.27	2.20*	.032	1.11
2	<i>L2 Proficiency</i>	.50	.224	.02	.009	.34	2.72*	.009	1.11
	<i>L1-RS</i>			1.30	.579	.28	2.24*	.029	1.11
3	<i>1st Exposure</i>	.53	.241	-.18	.193	-.28	-.93	.355	1.03
	<i>L2 Proficiency</i>			.03	.008	.36	2.93*	.005	1.06
	<i>L2-RS</i>			1.94	.756	.31	2.56*	.013	1.03
4	<i>L2 Proficiency</i>	.52	.243	.03	.008	.38	3.15*	.003	1.03
	<i>L2-RS</i>			1.92	.755	.30	2.54*	.013	1.03

* 1st Exposure (Age of first exposure to L2)

Hierarchical multiple regressions were also conducted to test whether there was a relationship between the dependent variable (RC scores) and the predictor variables (L1-RS and L2-RS) after controlling for the impact of L2 proficiency. As shown in Table 4, L2 proficiency alone (Model 1) significantly predicted L2 RC performance, $F(1, 54) = 11.98, p = .001$, Adjusted $R^2 = .16.6$. That is, 16.6% of the variance in L2 RC performance could be predicted by knowing the students' proficiency. Model 2 showed that L2-RS had significant predictive power in L2 RC even after controlling the effects of proficiency, $F(1, 53) = 5.03, p = .029$, Adjusted $R^2 = .22.4$, with a medium effect. That is, the L1-RS explains 5.8% additional variance, R^2

change = .058. The beta weights suggest that L1-RS contributes significantly to predicting L2 RC performance. Model 3 showed that L2-RS also had significant predictive power in L2 RC after controlling the effects of proficiency, $F(1, 53) = 6.46$, $p = .014$, Adjusted $R^2 = .243$, with a medium effect. That is, L2-RS explains 7.7% additional variance, R^2 change = .077. VIF was not found in any multivariate analysis. The findings corroborate previous studies (e.g., Coughlin and Tremblay, 2010; Harrington and Sawyer, 1992; Miyake and Friedman, 1998) in that L2-RS had a stronger predictive power for L2 RC performance than does L1-RS.

Table 4. Model summary of L1-RS and L2-RS: Predicting RC performance

<i>Model</i>	<i>Predictor</i>	<i>R</i>	<i>Adjusted R²</i>	<i>B</i>	<i>SEB</i>	<i>β</i>	<i>t</i>	<i>p</i>	<i>VIF</i>
<i>1</i>	<i>L2 Proficiency</i>	.43	.166	.03	.009	.43	3.46*	.001	1.000
<i>2</i>	<i>L2 Proficiency</i>	.50	.224	.02	.009	.34	2.72*	.009	1.106
	<i>L1-RS</i>			1.30	.579	.28	2.24*	.029	1.106
<i>3</i>	<i>L2 Proficiency</i>	.52	.243	.03	.008	.38	3.15*	.003	1.029
	<i>L2-RS</i>			1.92	.755	.30	2.54*	.014	1.029

The findings suggest that the students' sufficient knowledge in the L1 allows them to easily go through the continuous processing of input in L1 sentences. Conversely, their lack of automaticity in L2 reading consumes extra WM resources, which in turn results in a reduction in capacity (e.g., McLaughlin 1995; Miyake and Friedman 1998; Osaka and Osaka 1992; Osaka *et al.* 1993). That is, the not fully developed L2 proficiency of students impedes them from drawing the same WM resources when they process L1 texts in L2. The results also support previous studies (e.g., Coughlin and Tremblay 2010; Harrington and Sawyer 1992; Miyake and Friedman 1998; Walter 2004, 2007) in that L2-RS better reflected less efficient processing and difficulty restoring information in L2 reading.

3.2 Experiment 2

Experiment 2 further examines the role of RS in L2 RC by taking into account two reading dimensions, factual and inferential understanding. It is assumed that the effects of RS can depend on the complexity of each level of comprehension (e.g., Engle and

Marshall 1983; Just and Carpenter 1992; Kintsch *et al.* 1990; Rosen and Engle 1997). In addition, it also investigates the relationship between RS and retention of information in L2 texts. Based on Experiment 1 that L2-RS rather than L1-RS better predicted the students' L2 RC performance, L2-RS was used for group classification. The following are specific research questions addressed in the second experiment:

1. Are there any differences between high- and low-reading span groups in L2 RC scores on the factual and inferential questions?
2. Are there any differences between high- and low-reading span groups in the amount of changes that occur over time in scores on two different memory tests?

3.2.1 Methodology

3.2.1.1 Participants

Fifty Korean university students (*Male* = 26, *Female* = 24, *Mean age* = 24.28 years) at a similar proficiency level to those in Experiment 1 were recruited and were paid for their participation for the study.

3.2.1.2 Materials

Experiment 2 is based on the same material used in Experiment 1, including the language background information questionnaire and L2-RS test. In addition, two other materials were also included: a RC test based on the two-level comprehension (factual and inferential) and a self-rating questionnaire asking about their previous knowledge on the texts.

3.2.1.2.1 Prior knowledge questionnaires

A prior knowledge questionnaire was constructed in order to ensure that all students had a similar degree of background knowledge about the topics in the texts. It is assumed that prior knowledge can affect the effective processing of the contents. A 5-point Likert scale was used, with 1 being never heard of the topic

before and 5 being very familiar with the topic.

3.2.1.2.2 Reading comprehension test

Fifteen texts were selected for a RC test through a pilot study. Readability of the texts were controlled by using a similar number of words in each text (mean words per text = 120). The texts consisted of various topics, including animals, culture, history, and events. The RC test for immediate and delayed memory tests consisted of thirty multiple-choice questions. Each text contained two types of questions, comprising factual (textbase) and inferential (situational model) comprehension.

Experiment 2 employed timed reading assumed to better explain the effects of L2-RS on L2 RC performance, especially in high-level processing. Reading time for each text was set through a pilot study. The longest time was selected as reading time for each text in order to provide all the students with sufficient time to process the texts. For the delayed memory test, the same questions were used in a different order to eliminate order effects.

3.2.1.3 Procedure

Each student was asked to attend two sessions for the study: a sixty-minute session for the L2-RS test and computer-based RC test (immediate memory test), and a ten-minute session for the delayed memory test. The procedure of the first session was the same as Experiment 1. Right after the RC test, the students were instructed to rate how familiar they were with each of the fifteen topics by using a 5-point Likert scale. The delayed memory test was held the following day after the RC test, and the students were instructed to give answers based on their memory by presenting the same questions in a different order.

3.2.1.4 Scoring

The students were placed into either a high-RS or low-RS group based on their RS scores. The RS test and RC test followed the same scoring system used in Experiment 1. The students' level of familiarity corresponded to their number of choice. For instance, if the students chose 5, they knew and were very familiar with

the topic of the text.

3.2.1.5 Data analysis

Only forty-four students were included for the data analysis since six students, who did not attend the two sessions, were excluded. There was no significant difference between the high-RS ($M = 2.18$) and low-RS ($M = 2.44$) groups in terms of background knowledge of the topics contained in the texts: $t(43) = 1.416$, $p = .164$.

A $2 \times 2 \times 2$ three-way mixed univariate Analysis of Covariance (ANCOVA) was conducted to assess whether there were group, question type, and time differences in the comprehension scores: two within-subjects factors (Time and Question Type) and one between-subjects factor (Group). The participants' TOEIC scores served as a covariate since TOEIC scores may be a confounding or interacting variable. Alpha level was set at .05 for all the statistical analysis.

3.2.2 Results and discussion

3.2.2.1 Descriptive statistics

The assumptions of normality of all the variables used for Experiment 2 were met in that skewness and kurtosis were within ± 2 . A t -test showed that there were no significant differences in TOEIC scores between the high-RS ($M = 838.64$, $SD = 59.00$) and low-RS ($M = 812.05$, $SD = 56.58$) groups, $t(42) = 1.526$, $p = .135$. In addition, the correlations between the covariate and the types of comprehension questions were significant on the immediate memory test: factual ($r = .372$, $p = .013$) and inferential ($r = .327$, $p = .030$), respectively. There were also significant correlations between the covariate and the two-type comprehension questions on the delayed memory test, factual ($r = .367$, $p = .014$) and inferential ($r = .350$, $p = .037$), respectively. Accordingly, TOEIC scores are appropriate to use as a covariate.

Descriptive statistics of variables are summarized in Table 5. In general, the students received higher scores on the factual than on the inferential questions in the immediate memory test, $M = 11.73$ and $M = 9.11$, respectively. They also received higher scores on the factual than on the inferential questions in the delayed memory

test, $M = 10.93$ and $M = 8.77$, respectively.

Table 5. Descriptive statistics ($N = 44$)

Memory Test		Group	M	SE	Memory Test		Group	M	SE
Factual	Immediate	High	12.64	0.26	Inferen- tial	Immediate	High	9.97	0.42
		Low	10.81	Low			8.25		
	Delayed	High	12.04			0.30	Delayed	High	9.56
		Low	9.83	Low		7.98			
	Total	High	12.34			0.26	Total	High	9.77
		Low	10.32	Low		8.12			

*covariate = TOEIC scores, SE =Standard error High-RS ($N = 22$), Low-RS ($N = 22$)

3.2.2.2 Main effects of group, question type, and time on comprehension scores

Box's test showed the assumption of homogeneity of covariance matrices was met. In addition, the assumptions of linearity and uncorrelated errors were checked and met.

3.2.2.2.1 The effect of group

As shown in Table 5, the high-RS group outperformed the low-RS group on the RC tests containing factual ($M = 12.34$ and $M = 10.32$, respectively) and inferential ($M = 9.77$ and $M = 8.12$, respectively) questions. In addition, the high-RS group received consistently higher scores in both the immediate and the delayed memory tests. Specifically, in the immediate memory test, the high-RS group received higher scores than their counterparts on both the factual ($M = 12.64$ and $M = 10.81$, respectively) and inferential ($M = 9.97$ and $M = 8.25$, respectively) understanding. The same results were found in the delayed memory test. The results of the between-subjects ANCOVA revealed a significant main effect of Group, $F(1, 41) = 28.28$, $p < .001$, partial eta squared = .42, as shown in Table 6.

The findings support previous studies (Alptekin and Erçetin 2009; Conway 1996; Estevez and Calvo 2000; Service *et al.* 2002) in that RS is an important factor that affects both factual and inferential comprehension in L2 reading. The findings suggest that the high-RS students can effectively allocate resources for processing the

information in the L2 texts. They are also better at storing the products of the processes more readily, which in turn facilitates retrieval and memory strengthening.

3.2.2.2.2 The effect of question type

Although the students were better at understanding factual than inferential questions, $M = 11.33$ and $M = 8.94$, respectively. The test of within-subjects effects showed no significant differences between the types of comprehension questions, $F(1, 41) = .66$, $p = .422$. The findings contradict previous studies (e.g., Alptekin and Erçetin 2009). The findings suggest that inferential comprehension is more demanding and involves “a series of complex tasks associated with implicit information” in a text (Alptekin and Erçetin 2009: 635). Competent L2 readers have the ability to make “more native-like selection of cues, avoiding irrelevant and conflicting information, and decreasing WM load” (Service *et al.* 2002: 406). However, the students in the experiment had not yet developed to a point of automaticity in L2 reading. Furthermore, they were unfamiliar with L2 rhetorical structures and lacked linguistic knowledge. Accordingly, the students might have allocated a significant amount of time and resources in shallow processing, such as factual understanding (e.g., Kintsch *et al.* 1990; Perkins and Bratten 1992).

3.2.2.2.3 The effect of time

The changes from the immediate to delayed memory tests on the RC scores were small, $M = 10.42$ and $M = 9.85$, respectively. The tests of within-subjects effects showed no significant main effect of time on the memory tests, $F(1, 41) = .50$, $p = .485$. The findings contradict Kintsch *et al.*'s (1990) L1 study where a significant effect of time was observed at the textbase-level (factual) comprehension. One plausible explanation for the weak relationships between RS and retention of memory is the within-subjects design. For example, the same memory tests (immediate and delayed) might have provided the students with the opportunity to consolidate their understanding. Another explanation is that the time interval between the memory tests might not be long enough to measure memory decay.

3.2.2.2.4 Interaction between group, question type, and time

The results showed that there were no significant interactions between and among variables: Group x Time ($F(1, 41) = .32, p = .578$), Group x Question Type ($F(1, 41) = .28, p = .601$), Question Type x Time ($F(1, 41) = .09, p = .765$), and Group x Time x Question Type ($F(1, 41) = 1.14, p = .291$). The findings indicate that RS does play an important role in L2 RC performance.

Table 6. Effects of group and time on factual and inferential comprehension scores

<i>Source</i>	<i>df</i>	<i>F</i>	<i>p</i>	<i>Partial eta squared</i>
Group	1	28.28**	.000	.420
Q Type	1	.66	.422	.016
Time	1	.50	.485	.601
Group x Time	1	.32	.578	.008
Group x Q Type	1	.28	.601	.007
Q Type x Time	1	.09	.765	.002
Group x Time x Q Type	1	1.14	.291	.027
Error	41			

*covariate = TOEIC scores, Q type = Question Type

4. General discussion and conclusion

The present study conducted two experiments in order to explain the relationship of RS to L2 RC. Specifically, Experiment 1 explored the relative effects of L1-RS and L2-RS on overall L2 RC performance. Experiment 2 further examined the extent to which RS influences L2 RC and retention in terms of two different types of comprehension (factual and inferential).

First, L1-RS and L2-RS had a positive correlation with L2 RC, and L2-RS rather than L1-RS had a stronger predictive power for L2 RC performance. The findings indicate that L2-RS better represents the students' L2 processing in L2 reading. The stronger association between L2-RS and L2 RC performance can be attributed to the students' less automatic L2 processing ability than that of their L1.

Second, RS had a great impact on L2 RC performance. For instance, having a greater RS would provide the students with more resources to engage a deeper-level processing. It would also facilitate transfer of information into LTM for later use. Consequently, those with a greater RS outperformed those with a small RS on the RC tests measuring retrieval and retention.

Third, although the students were better at processing and retaining the factual than inferential questions, there were no significant effects of the question type on the L2 RC performance. Nonetheless, it is reasonable to assume that making inferences while reading is more challenging for L2 learners both those in the high-RS and low-RS alike. Unlike with their L1 processing, L2 learners use a substantial amount of WM to facilitate the reading task due to their not yet fully automatized L2 processing. Accordingly, they leave a limited amount of resources to generate inferences. Their poor inferential understanding is the result of diminished processing resources and inefficient strategies.

Lastly, there was no time effect on retaining the information in the L2 text. The contradicting results might have been influenced by the within-subjects design and short intervals between the two memory tests. In addition, there were no interactions between variables.

There are several limitations that merit discussion. First, the results of the study should be generalized with caution since the participants were university students at a similar level of English proficiency in Korea. Second, this study used a rather small number of participants. Finally, a number of variables, including anxiety, strategy, and linguistic knowledge, bring individual differences in L2 RC performance. Therefore, further research is necessary to incorporate multiple measures of individual differences. Despite the limitations, the present study shed some light on understanding of L2 RC. Specifically, it attempted to explain individual differences in L2 reading (discourse-level comprehension) performance in relation to RS. It also investigated the relationship between L1-RS and L2-RS.

To conclude, L2 learners are not flexible readers who possess the knowledge, skills, and strategies to effectively accommodate the demand of L2 processing. Therefore, it is important that L2 teachers know a potential variable, namely RS that can influence the language learning of their students.

References

- Alptekin, Cem and Gülcan Erçetin. 2009. Assessing the relationship of working memory to L2 reading: Does the nature of comprehension process and reading span task make a difference? *System* 37: 627-639.
- Alptekin, Cem and Gülcan Erçetin. 2010. The role of L1 and L2 working memory in literal and inferential comprehension in L2 reading. *Journal of Research in Reading* 33: 206-219.
- Baddeley, Alan. 2000. The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences* 4: 417-423.
- Baddeley, Alan. 2003. Working memory and language: An overview. *Journal of Communication Disorders* 26(3): 189-208.
- Burton, Christine and Meredyth Daneman. 2007. Compensating for a limited working memory capacity during reading: Evidence from eye movements. *Reading Psychology* 28: 163-186.
- Carroll, David. 2008. *Psychology of language (5th edition)*. London: Thomson Wadsworth.
- Cohen, Jacob. 1988. *Statistical power analysis for the behavioral sciences (2nd edition)*. Hillsdale, NJ: Erlbaum.
- Conway, Andrew. 1996. Individual differences in working memory capacity: More evidence for a general capacity theory. *Memory* 4: 577-590.
- Coughlin, Caitlin and Annie Tremblay. 2013. Proficiency and working memory based explanations for nonnative speakers' sensitivity to agreement in sentence processing. *Applied Psycholinguistics* 34: 615-646.
- Daneman, Meredyth and Patricia Carpenter. 1980. Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior* 19(4): 450-466.
- Daneman, Meredyth and Philip Merikle. 1996. Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin and Review* 3: 422-433.
- Dijkstra, Ton. 2005. Bilingual visual word recognition and lexical access. In Judith F. Kroll and Annette M. B. De Groot (eds.), *Handbook of bilingualism: Psycholinguistic approaches*, 179-201. New York: Oxford University Press.
- Engle, Randall and Kathy Marshall. 1983. Do developmental changes in digit span result from acquisition strategies? *Journal of Experimental Child Psychology* 36(3): 429-436.
- Estevez, Adelina and Manuel Calvo. 2000. Working memory capacity and time course of predictive inferences. *Memory* 8: 51-61.
- Felser, Claudia and Leah Roberts. 2007. Processing *wh*-dependencies in a second language: A cross-modal priming study. *Second Language Research* 23(1): 9-36.
- Gass, Susan and Iunkyu Lee. 2011. Working memory capacity, inhibitory control, and proficiency in a second language. In Monika Schmid and Wander Lowie (eds.), *Modeling bi-*

- lingualism: From structure to chaos: In honor of Kees de Bot*, 59-84. Amsterdam: John Benjamins Publishing.
- Grady, William O. 1991. *Categories and case: The sentence structure of Korean*. Amsterdam: John Benjamins Publishing.
- Harrington, Michael and Mark Sawyer. 1992. L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition* 14: 25-38.
- Juffs, Alan. 2004. Representation, processing and working memory in a second language. *Transactions of the Philological Society* 102(2): 199-225.
- Just, Marcel and Patricia Carpenter. 1992. A capacity theory of comprehension: Individual differences in working memory. *Psychological Review* 99: 122-149.
- Kim, Kitaek. 2015. Sensitivity to ungrammatical object drop in English by Korean young and adult L2 learners. *Linguistic Research* 32: 57-81.
- Kintsch, Walter. 1990. Sentence memory: A theoretical analysis. *Journal of Memory and Language* 29: 133-159.
- Kintsch, Walter, David Welsch, Franz Schmalhofer, and Susan Zimny. 1990. Sentence memory: A theoretical analysis. *Journal of Memory and Language* 29: 133-159.
- McLaughlin, Barry. 1995. Aptitude from an information processing perspective. *Language Testing* 11: 364-381.
- Miyake, Akira and Naomi Friedman. 1998. Individual differences in second language proficiency: Working memory as "language aptitude." In Alice F. Healy and Lyle E. Bourne (eds.), *Foreign language learning: Psycholinguistic studies on training and retention*, 339-364. Mahwah, NJ: Lawrence Erlbaum Associates.
- Ortega, Lourdes. 2009. *Understanding second language acquisition*. London: Hodder Education.
- Osaka, Mariko and Naoyuki Osaka. 1992. Language-independent working memory as measured by Japanese and English reading span tests. *Bulletin of the Psychonomic Society* 30: 287-289.
- Osaka, Mariko, Naoyuki Osaka, and Rudolf Groner. 1993. Language-independent working memory Evidence from German and French reading span tests. *Bulletin of the Psychonomic Society* 31(2): 117-118.
- Pae, Hye. 2011. Is Korean a syllabic alphabet or an alphabetic syllabary? *Writing Systems Research* 3(2): 103-115.
- Park, Hyangsook. 2015. *The relationship between reading span and L2 reading comprehension*. PhD Dissertation, Kyungpook National University.
- Perkins, Kyle and Sheila Brutten. 1992. The effect of processing depth on ESL reading comprehension. *Journal of Research in Reading* 15: 67-81.
- Rosen, Virginia and Randall Engle. 1997. The role of working memory capacity in retrieval. *Journal of Experimental Psychology: General* 126: 211-227.
- Sagarra, Nuria. 2007. Online processing of gender agreement in low proficient

- English-Spanish late bilinguals. In José Camacho, Nydia Flores-Ferrán, Liliana Sánchez, Vivian Deprez, and María José Cabrera (eds.), *Romance linguistics 2006: Selected papers from the 36th Linguistic Symposium on Romance Languages*, 240-253. Amsterdam: John Benjamins.
- Sagarra, Nuria and Julia Herschensohn. 2010. The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. *Lingua* 120: 2022-2039.
- Van den Noort, Maurits, Peggy Bosch, and Kenneth Hugdahl. 2006. Foreign language proficiency and working memory capacity. *European Psychologist* 11: 289-296.
- Service, Elisabet, Simola Marjut, Oili Metsänheimo, and Sini Maury. 2002. Bilingual working memory span is affected by language skill. *European Journal of Cognitive Psychology* 14(3): 383-408.
- Walter, Catherine. 2004. Transfer of reading comprehension skills to L2 is linked to mental representations of text and to L2 working memory. *Applied Linguistics* 25: 315-339.
- Waters, Gloria and David Caplan. 1996. The capacity theory of sentence comprehension: Critique of Just and Carpenter 1992. *Psychological Review* 4: 761-772.
- Wen, Zhisheng. 2014. Theorizing and measuring working memory in first and second language research. *Language Teaching* 47: 174-190.

Appendix

Sample sentences of reading-span tests

L1-RS	L2-RS
Level 2 - 결과로 한 사람의 능력을 판단하는 것은 바람직하지 못하다. - 가족의 행복을 지켜주는 힘은 서로에 대한 이해와 믿음이다. Level 3 - 식중독을 예방하는 방법 중 하나는 손을 깨끗이 씻는 습관이다. - 태풍이 우리가 오랫동안 준비하고 기다렸던 소풍계획을 망쳤다. - 발목까지 오는 긴 치마를 입고 있어서 그녀의 상처를 볼 수 없었다. Level 4 - 옥상에 심은 상추를 따서 어머니는 저녁식탁 위에 올려놓았다. - 화가들은 일반인들 보다 사물을 보고 관찰하는 능력이 뛰어나다. - 판매율을 높이기 위해서 그가 노력해야 하는 것은 서비스이다. - 세월이 흘러간다 해도 나는 지금 이 순간을 잊지 못할 것이다. Level 5 - 연주회에서 그는 너무 피곤하여 그만 곤하게 잠이 들었다. - 도시락을 책상 위에 올려놓고 그는 깊은 생각에 잠겨 있었다. - 풍요가 당신의 행복을 보장해 주지 않는다는 것을 알아야 한다. - 미래에 대한 막연한 불안이 그를 매일 밤 잠 못 이루게 한다.	Level 2 - I am not hungry because I had some noodles. - During this weekend, I will work in the garden. Level 3 - She cannot go to school today because of flu. - They stopped playing football because of rain. - She gave me some help on my math task. Level 4 - She was supposed to bring the story book. - She opened the heavy wooden door in the cabin. - The woman dressed in blue is my favorite aunt. - I found it difficult to finish the work in two hours. Level 5 - The man is the best trained person for the project. - She really wanted to marry the man of her dream. - All the evidence suggests that he stole the money. - Play is essential for animals to develop social skills.

- 고아들이 소외되지 않고 보살핌을 받도록 사회가 보장해야 한다. Level 6 - 실수로 일어난 사고는 고의적인 행동과 구분되어야 한다. - 박물관의 전시품 중에서 우리에게 가장 인상적인 것은 가면이다. - 소나기가 갑자기 내려서 소년들은 나무 밑으로 급하게 달려갔다. - 딸기를 소금에 씻으면 살균효과도 있고 단 맛이 더 강해진다. - 달력을 보고서야 그녀는 오늘이 여동생의 생일인 것을 알았다. - 교실에서 장난을 하다가 한 학생이 꽃병을 떨어뜨리고 말았다.	- She loves to chat with her friends over the phone. Level 6 - People arrived to see the famous vocal band group. - They continued their trip despite the cold weather. - He developed his housing project within the year. - When he was young, he used to walk to school. - Many people joined in creating the modern bike. - Using the right brush is important for your teeth.
--	--

Hyangsook Park

Department of English Education
 Kyungpook National University
 80 Daehak-ro, Buk-gu, Daegu 41566, Korea
 E-mail: suziekim2001@gmail.com

Kichun Nam

Department of Psychology
 Korea University
 145 Anam-ro, Seongbuk-gu, Seoul 02841, Korea
 E-mail: kichun@korea.ac.kr

Yae-Sheik Lee

Department of English Education
 Kyungpook National University
 80 Daehak-ro, Buk-gu, Daegu 41566, Korea
 E-mail: yaesheik@knu.ac.kr

Received: 2016. 06. 24.

Revised: 2016. 09. 05.

Accepted: 2016. 09. 05.