Phonological Short-term Memory's Contribution to the L2 Reading Proficiency of Japanese EFL Learners

Akiko Kondo
Hyogo University of Teacher Education

This study investigated whether phonological short-term memory (PSTM) capacity has a significant relationship with the reading proficiency of Japanese English as a Foreign Language (EFL) learners, and the degree to which PSTM capacity contributes to L2 reading proficiency. For this purpose, the PSTM of 208 Japanese university students majoring in education and engineering was measured using an L1-based digit span test and an L1-based pseudoword span test, and reading proficiency was examined with a reading section of a standardized English proficiency test (Visualizing English Language Competency Test). The results of the regression analyses revealed that PSTM had significant positive effects on L2 reading, including its subcomponents. The study demonstrates the positive influence of PSTM on L2 reading proficiency, which previous studies have failed to do and provides insight into our understanding of the effects of PSTM on L2 reading proficiency.

本研究は、日本人英語学習者のリーディング力に、言語適性の一つである音韻的短期記憶力が与えている影響を調査することを目的に実施した。工学と教育学を専攻とする日本人大学生208名を対象に実験を実施した。実験参加者の音韻的短期記憶力は、日本語の数字暗証課題と非単語暗唱課題で、そして英語のリーディング力は語彙、文法、読解のセクションで構成されるVELCテストで測定した。各テストスコアを回帰分析によって検査したところ、音韻的短期記憶力は日本人英語学習者の英語のリーディング力に有意な影響を与えていることが示された。本研究は、先行研究ではまだ十分に調査がなされていない外国語のリーディング力の個人差要因としての音韻的短期記憶力の影響の理解に貢献できるものである。

https://doi.org/10.37546/JALTJJ43.2-2

JALT Journal, Vol. 43, No. 2, November 2021

Keywords*:* cognitive factor; grammar; individual differences; regression analyses; second language acquisition; text comprehension; vocabulary

Japanese EFL learners and teachers mostly focus on improving reading skills in the classroom due to their importance during university entrance examinations (MEXT, 2018; Watanabe, 2018). However, less attention has been given to individual differences (ID) regarding cognitive factors affecting L2 reading proficiency in a Japanese EFL context. Phonological short-term memory (PSTM) could be one of the cognitive ID factors that can explain L2 reading proficiency. PSTM is a storage subcomponent of working memory (WM), along with the central executive (an attention control system responsible for integrating information from other subcomponents and long-term memory), visuospatial sketchpad (storage subcomponent that handles visual images and spatial information), and episodic buffer (storage subcomponent that is involved in episodic representation) (Baddeley, 2010). The PSTM, which Baddeley and Hitch (1974) call the *phonological loop*, is a language user's capacity to temporally hold sound information, including both verbal and acoustic elements of speech (Baddeley, 2000, 2003; Baddeley & Hitch, 1974). This storage component consists of two subcomponents: a temporary storage system and a subvocal rehearsal system. The storage system "holds memory traces over a matter of seconds, during which they decay, unless refreshed by the subvocal rehearsal system" (Baddeley, 2003, p. 191). The subvocal rehearsal system, where participants subvocalize the items to be memorized, and which maintains the information within the store also serves "the function of registering visual information within the store, provided the items can be named" (Baddeley, 2003, p. 191). In regard to the processes the storage and subvocal rehearsal system are responsible for, the PSTM can be assumed to involve the reading process, because when people read, they need to perceive verbal items visually presented by letters, name the items, and subvocalize the items in their mind to interpret their meaning. PSTM enables people to read through a process where they maintain words by activating their phonological representation to create the meaning of the text. Therefore, individual differences in the capacity of PSTM might explain variance in learners' reading skills.

PSTM and Fundamental L2 Components

Vocabulary and Grammar

Before reviewing previous research on PSTM and L2 reading skills, the effects of PSTM on L2 vocabulary and grammar knowledge are briefly discussed because vocabulary and grammatical knowledge are fundamental components of reading. The relationship between PSTM and L2 vocabulary acquisition has been extensively studied, and numerous researchers have demonstrated a direct link between PSTM and L2 vocabulary development (e.g., Atkins & Baddeley, 1998; Martin & Ellis, 2012; Masoura & Gathercole, 1999; Service, 1992; Service & Kohonen, 1995). For example, Masoura and Gathercole found a significant correlation (r = .36) between L2 vocabulary development and L1 PSTM in 8- to 11-year-old Greek children. The unique feature of their study is that the researchers assessed both PSTM and vocabulary knowledge in two languages, Greek (L1) and English (L2), enabling them to make direct comparisons of the strength of the association across languages. L2 vocabulary knowledge was significantly correlated with not only L1 PSTM (r = .36), but also L2 PSTM (r = .39). In addition, L1 vocabulary knowledge was significantly correlated with L1 PSTM (r = .50) and L2 PSTM (r = .35).

Further, researchers have found that PSTM significantly influences L2 grammar knowledge as well (French & O'Brien, 2008; Martin & Ellis, 2012; O'Brien et al., 2006). For example, O'Brien et al. found PSTM played an important role in the L2 grammar knowledge of 43 adult L2 learners of Spanish; PSTM measured by a pseudoword recognition task explained 5.4% of the variance in the correct use of function words. O'Brien et al. also examined the contribution of PSTM to L2 grammar knowledge across proficiency levels and revealed that PSTM explained a significant amount of variance (15.7%) in the correct use of function words with high-proficiency participants but not with low-proficiency participants. O'Brien et al. stated that while at earlier stages of L2 learning, low-proficiency learners concentrate on using content words and use PSTM for lexical access, while in the later stages of L2 learning, high-proficiency learners use PSTM to learn more complex grammatical forms, as lexical access is easier and therefore places less burden on the memory system.

Further evidence for the effects of PSTM on L2 grammar was reported by a study comprising a larger group of young L2 learners (French & O'Brien, 2008). The role of PSTM in L2 grammar knowledge was examined in 104 elementary school English learners (M = 11 years old), and PSTM capacity—as measured by two non-word repetition tests—was found to explain almost 30% of the variance in L2 grammar knowledge at the end of instruction. The study's significant features are its focus on the effects of PSTM on L2 grammar gains rather than grammatical knowledge at the start of the experiment and the fact that the variance attributed to intelligence and prior

L2 knowledge were partialled out with a hierarchical regression analysis. The study also demonstrated that PSTM in L2 learning can improve as learners are exposed to aural input from the target language.

Furthermore, the effects of PSTM on grammar knowledge have been demonstrated to be connected to the previous knowledge of the language learner. For instance, Martin and Ellis (2012) investigated the role of PSTM in learning the grammar and vocabulary of an artificial language among 40 monolingual English speakers recruited from a large American university. Two hierarchical multiple regression analyses revealed that the L1 English speaking adults' (N = 40) PSTM capacity, tapped by non-word recognition tests, explained 17% and 10% of the variance in the L2 receptive and productive grammar test scores, respectively. As their approach used an artificial language as the target language, the researchers were able to examine the link between PSTM and learning grammar independent of previous linguistic knowledge of the target language.

PSTM and L2 Reading

The review of previous related studies has shown the association of PSTM and the acquisition of L2 vocabulary and grammar knowledge (including an artificial language) across young and adult learners. Given that L2 reading requires learners to process text by engaging their L2 vocabulary and grammar knowledge, the previous research suggests PSTM may be positively associated with L2 reading processing. Moreover, as discussed in the introduction, while reading, people subvocalize the visually presented words of a sentence and hold the information for interpreting the meaning making, which involves PSTM.

However, several studies failed to show effects of PSTM on L2 reading. Harrington and Sawyer (1992), one of the most widely cited studies on the effects of PSTM on L2 reading, did not find a significant correlation between PSTM (measured by L2-based digit and word span) and L2 reading in 34 Japanese university students who were advanced EFL learners. The study did find, however, a strong correlation with WM as measured with the L2-based reading span test. In addition to the relatively small number of participants, another issue with this study was its use of an L2-based memory span test. PSTM measured with L2-based tests can be highly influenced by L2 proficiency. Hummel and French (2010) also pointed out that Harrington and Sawyer's null results (the non-significant correlation between PSTM and L2 reading skills) might have been due to the fact that they did not consider the possibility of language or lexicality effects on memory span tests (the involvement of the L2 proficiency caused by using the target language for measuring their PSTM), including the L2-based digit, word, and reading span tests. Furthermore, as their participants were advanced EFL learners, it is uncertain if their results can be applied to intermediate or lower-level language learners, as the effects of PSTM on learning L2 vocabulary tend to be smaller in higher proficiency learners (Cheung, 1996; French, 2006; Hummel, 2009).

Also working with advanced learners, Hummel and French (2016) showed that PSTM could predict the L2 reading proficiency of 45 French speaking L2 learners. One of the major differences between this study and Harrington and Sawyer (1992) is the language used for measuring PSTM. Hummel and French measured PSTM using Arabic-based non-word repetition, with Arabic being an unfamiliar language to the participants, and L1 French-based serial recognition tasks. They controlled the language effects by avoiding using the L2 in the measurement of PSTM. Using regression analyses, Hummel and French demonstrated the predictability of PSTM on L2 reading proficiency. However, one methodological limitation of their study is its small sample size. The number of participants (N = 45) is insufficient for regression analyses. According to Tabachnick and Fidell (2007), a rule of thumb for conducting the analyses is "$N \geq 104 + m$ for testing individual predictors" (p. 123).

The methodological issues in previous studies, such as small sample sizes and the languages used in the PSTM measure, were addressed by Kormos and Sáfár (2008) in another frequently cited study. Their study involving 121 Hungarian secondary school students (15-16 years old) supported the results of Harrington and Sawyer (1992). Their analysis revealed no significant correlation between PSTM capacity as measured with the participants' L1-based non-word repetition test and L2 reading scores from the Cambridge First Certificate Exam, which was found for both beginning (n = 100) and intermediate L2 learners (n = 21).

In addition to Hummel and French (2016), the positive influence of PSTM on L2 reading skills was demonstrated by Swanson et al. (2011), who found significant effects for PSTM on L2 reading skills in 471 Hispanic elementary school children in the United States. In Swanson et al.'s study, the participants' PSTM span was measured by L1 Spanish-based forward digit, backward digit, word, and nonword span tests as one latent variable. A hierarchical regression analysis demonstrated that PSTM was a significant predictor of L2 (English) reading skills.

Methodological Issues in Previous Studies

Although studies have shown a significant influence of PSTM on L2 knowledge and skills, there are methodological issues that are worth discussing. When measuring participants' PSTM with recall or repetition tests, some researchers asked participants to recall L2-based items in memory span tests (Harrington & Sawyer, 1992; Nakanishi, 2011). For example, in the digit span test for Japanese EFL learners, researchers asked participants to recall digits in English, instead of Japanese. However, several studies pointed out a multicollinearity problem caused by using the same target language when measuring L2 memory and L2 proficiency (French & O'Brien, 2008; Hummel, 2009; van den Noort et al., 2006). For example, French and O'Brien measured native French-speaking participants' PSTM using English-based and Arabic-based non-word repetition tests at the outset and end of five months of intensive English instruction. The participants' PSTM as measured by an English-based non-word repetition test improved with their English development, while PSTM as measured by the Arabic-based non-word repetition test remained stable. This result also implies that language proficiency significantly affects PSTM span. Considering the results of these studies, it is preferable to use the participants' L1 when measuring PSTM to avoid the influence of their target language proficiency. In addition to the language used, the type of task also needs to be considered. Some studies include a manipulative process in measuring PSTM. For example, Swanson et al. (2011) included the backward digit span test as one of the PSTM measures. However, the test requires participants to orally produce digits backwards, which involves an additional manipulative memory process in which they need to reorder the digits they heard before oral reproduction. In fact, this type of test has been conventionally used as an instrument to measure WM capacity (Gathercole et al., 2004) and is therefore not suitable for PSTM. Thus, the instrument should not include a manipulative memory process.

Some research findings suggest that the degree of the effects of PSTM varies depending on the students' L2 proficiency level (Cheung, 1996; French, 2006; Hummel, 2009; O'Brien et al., 2006). For example, Cheung (1996) indicated that the effects of PSTM decrease as L2 proficiency increases. His investigation of 84 Hong Kong seventh graders (12.2 years old on average) showed that PSTM measured by a non-word span test predicted success in learning new foreign language words but this relationship was significant only in students whose L2 vocabulary was smaller than average (15% of the variance explained). These previous studies indicate that the role of PSTM is greater for lower-proficiency learners, implying that the effects of PSTM interact with long-term L2 knowledge. In learning novel words, more proficient learners can make use of long-term L2 knowledge that less-proficient learners have to a more limited extent. Furthermore, O'Brien et al.'s (2006) results also indicated the influence of proficiency on the relationship between PSTM and grammar knowledge. They found that at earlier stages of L2 learning, low-proficiency learners concentrate on the use of content words and use PSTM for lexical access, whereas, in the later stages of L2 learning, high-proficiency learners use PSTM to learn more complex grammatical forms, as lexical access is easier and therefore places less of a burden on the memory system. As these studies on L2 vocabulary and grammar showed, PSTM capacity negatively affects lower-proficiency learners more than higher-proficiency learners. Harrington and Sawyer (1992) in fact failed to find any significant influence of PSTM on advanced L2 learners' reading skills.

As discussed, the methodological issues of these studies suggest that further investigation is needed before the field can reach an informed position on the influence of PSTM capacity on L2 reading. To address these gaps in the literature, this study aims to investigate effects of PSTM capacity on intermediate-proficiency Japanese EFL learners' L2 reading proficiency, including receptive vocabulary, grammar, and text comprehension.

Method

Participants

This study was carried out with the participation of 208 post-secondary students from two institutions. One institution is a technical college offering engineering education and the other is a university of teacher education. Both institutions are relatively small, national schools located in a suburban area of western Japan. All participants were L1 Japanese-speaking students (M = 19.9 years old) majoring in engineering or education. The reported number of years of prior English study was between 8 to 10 years. At the time when this study was conducted, the results of each school's placement test indicated the participants' proficiency was around A2-B1 on the CEFR. Each week they attended 1-2 hours of integrated English classes that were designed to improve their English language skills (including reading, listening, speaking, and writing).

Invitations to participate were distributed to English classrooms at the university and sent to students using the schools' e-mail system. Participation was not required, and participants received 5,000 Japanese yen for their cooperation. The purpose of the study, the tasks they would be asked to complete, the time required, and how the data would be published were explained. Students who understood the study and wished to participate signed up by accessing a free online scheduling service, *densuke* (https://www.densuke.biz/), which was used to recruit students and arrange the data collection schedule.

Instruments

Measuring L2 Reading Skills (VELC Test)

The English proficiency of the participants, who were streamed by the results of an entrance examination, demonstrated a narrow distribution that was too low for TOEFL or TOEIC, both of which target higher-proficiency test-takers. Therefore, the participants' L2 reading skills were measured using the Visualizing English Language Competency Test (VELC Test) designed by Kinseido, an English textbook publisher in Tokyo. This test was chosen as it could appropriately measure the proficiency range of the Japanese university students (Kumazawa et al., 2016).

As the item-level data were not provided by the testing company, the reliability of this test cannot be calculated. However, Shizuka and Mochizuki (2014), who are part of the group who developed the VELC Test, reported that the coefficient of reliability was high (Rasch person reliability = .95) and its multiple correlation coefficient to TOEIC scores was .82 based on a study of 5,583 Japanese university students. Furthermore, Kumazawa et al. (2016) provided evidence indicating that the VELC Test ($k = 120$) was reliable with a small margin of error based on a study with 4,407 Japanese university students.

The VELC Test consists of a listening section and a reading section, each of which includes three parts with 20 items in each part, totaling 60 listening items and 60 reading items. For the current analysis, only the VELC Test Reading section (VTR) scores were used. The VTR consists of three parts: Part 1 (vocabulary; VTR-Vocabulary), Part 2 (grammar: sentence structure; VTR-Grammar), and Part 3 (text comprehension; VTR-Comprehension). Examinees are given 45 minutes to complete the VTR.

In the VTR-Vocabulary, the participants chose one English word from four options that best corresponds to the meaning of a given Japanese word or set of words. Individual responses were hand written on an exam sheet. In the sample item below, the correct answer is "(B) experience" because it is the English word that conveys the same meaning as the Japanese word, *keiken*, *taiken*, (経験、体験). The VTR-Vocabulary test items were selected from the JACET 1000-7000 level vocabulary list (JACET Basic Word Revision Committee, 2003) to measure written receptive English vocabulary knowledge.

Sample item:
経験，体験　(A) society　*(B) experience　(C) notice　(D) language

In the VTR-Grammar part, the participants must complete an incomplete sentence by selecting one location to insert a target word from four options. In the sample below, the correct answer is (a) because the word *who* should be inserted in place of (a) for the sentence to be grammatically accurate.

Sample item:
Today, people *(a) can use the Internet (b) find it easy to (c) communicate with (d) each other.　[who]

In the last part, VTR-Comprehension, the participants read several English sentences that make up a coherent passage. One sentence contains a blank space in which examinees must choose a word or phrase from four options that will complete the sentence. In the sample below, (b) is the correct option to make the sentence meaningful within the context of the other sentences in the passage. The length of the passages varied from approximately 20 to 80 words, and the passages scored 30 to 80 on the Flesch Reading Ease index (Kumazawa, 2015). This part assesses the ability to understand the content of each sentence and the relationship between those sentences.

Sample item:
Service animals are not pets. People keep pets for fun and companionship. People keep service animals because they are ___________. A guide dog, for example, helps people who cannot see.
(A) beautiful　*(B) useful　(C) fun to play with　D) fun to look at

The VELC Test scoring adopts a standard procedure in which each participant's scores (both total and sub scores) are transformed so that the mean score is 500 and the standard deviation is 100; thus, test-takers know whether their score is higher or lower than the mean score of Japanese university students who participated in the pilot study for developing the

VELC Test. For example, a score of 550 indicates that the score is higher than the average Japanese university student by 0.5 times the standard deviation (Shizuka & Mochizuki, 2014).

Measuring PSTM

To measure the participants' PSTM capacity, two widely used PSTM tests were adopted. First, the forward digit span test was designed to assess the participants' capacity to memorize L1 digit information over a short period of time. In this computer-based test, which took 10 minutes to complete, the participants listened to a set of digits in their L1, Japanese, and orally reproduced the digits in the same order as they had heard them; their responses were then recorded on the computer. The test consisted of 16 items, each composed of 6 to 11 digits. The test structure, including the range of digit numbers and item numbers for each level, was determined by a series of pilot tests. Participants were scored between 0 and 10 per item based on the percentage of digits reproduced correctly. For example, if a participant said 3413698123 for the test item that required the reproduction of 3413698175, a score of 8.0 points was assigned for this item because eight of the ten digits (80%) were reproduced. Likewise, if a participant produced 3413679815 for the same target item, a score of 6.0 points was assigned because six of the ten digits (3, 4, 1, 3, 6, and 5) were reproduced in their original position in the item.

The pseudoword repetition test, the second test for measuring PSTM capacity, is also a computer-based test; it has the same format as the forward digit span test and also takes 10 minutes to complete. In this test, participants listened to and orally reproduced a set of pseudowords consisting of three different Japanese phonemes mora, which sound like Japanese words but do not have any meaning. In total, 66 different pseudowords were used for this test, each of which included three different Japanese phonemes. The test consisted of 17 items, with 1 two-pseudoword item, 6 three-pseudoword items, 8 four-pseudoword items, and 2 five-pseudoword items (see the sample items below). The test structure was examined for reliability and validity through a series of pilot tests and was found to be acceptable. For the scoring procedure, as with the forward digit span test, each item was scored between 0 and 10 based on the percentage of the pseudowords (calculated based on the number of syllables) correctly reproduced by the participant.

Sample item:
No.1 げのて (*ge-no-te*). みたじ (*mi-ta-ji*).
No.8 まこそ (*ma-ko-so*). できや (*de-ki-ya*). よみと (*yo-mi-to*). なおて (*na-o-te*).

The Cronbach's alpha was calculated using SPSS for the designed tests (the forward digit span test and pseudoword repetition test). The reliability estimate demonstrated good consistency (a =. 90) for each test.

Data Collection and Analysis

Each participant took all tests on one day chosen from several options. The session took approximately 2 hours in total, including the tutorial, which consists of an explanation of the purpose of the study, the procedures for each test, and how data would be kept confidential and reported anonymously.

Before performing the main analyses, the descriptive statistics of each instrument were checked, and data were screened to identify outliers. Next, four standard multiple regression models were employed to investigate the extent of variance in L2 reading skills that was explained by the PSTM variables. The first model considered the PSTM scores (the scores of the forward digit span test and pseudoword repetition test) as the independent variables and the L2 (English) reading skill level (the total scores of the VELC reading test) as the dependent variable. This was followed by three models wherein the independent variables were the scores of the forward digit span test and pseudoword repetition test, and the dependent variables were the scores of the VTR-Vocabulary, VTR-Grammar, and VTR-Comprehension tests, respectively.

Results

Descriptive Statistics

Table 1 provides descriptive statistics for the L2 reading variables (VTR total score, VTR-Vocabulary score for written receptive vocabulary, VTR-Grammar score for grammar, and VTR-Comprehension score for comprehension) and the predictor variables (scores on the forward digit span test and the pseudoword repetition test). Here, before proceeding with the main analyses, univariate and multivariate outliers were checked, and cases 36 and 75 on the forward digit span test were identified as univariate outliers based on the z-score criterion of ±3.29 (Tabachnick & Fidell, 2007); thus 206 cases were used in the main analysis. In addition, normality of distribution was checked by calculating the skewness and kurtosis statistics, standard errors, and *z*-scores. The results showed that the data for all instruments were normally distributed.

Table 1

Descriptive Statistics of the L2 Reading and PSTM Variables

			95% CI			
	M	*SD*	*LL*	*UL*	Min.	Max.
VTR-Total	489.42	71.47	479.60	499.24	346	673
VTR-Vocabulary	502.16	76.09	491.70	512.61	342	724
VTR-Grammar	499.80	73.23	489.74	509.86	354	677
VTR-Comprehension	456.23	79.48	464.32	486.15	259	649
Forward Digit Span Test	77.56	20.93	74.69	80.44	11	149
Pseudoword Repetition Test	85.85	21.56	82.89	88.81	32	147

Note. N = 206. PSTM = phonological short-term memory; VTR = VELC Test Reading section.

Results of the Regression Analyses

Table 2 shows the results of the four models of the multiple standard regression analyses. For total score of VTR, the results indicate that the model predicted variance that was significantly greater than zero, F (2, 203) = 6.05, p = .003, with R^2 at .056, and that the variance explained by PSTM capacity was 5.6%, which is small but statistically significant. Only PSTM measured by the forward digit span test significantly predicted L2 reading skills (β = .26, p = .001) as measured by the VTR, whereas the pseudoword repetition test did not contribute to explaining the significant variance in L2 reading skills (β = -.10, p = .205).

The results of the standard multiple regression analysis to investigate the amount of variance of L2 receptive vocabulary explained by PSTM capacity indicate that the model predicted variance that was significantly greater than zero, F (2, 203) = 5.54, p = .005, with R^2 at .052. The variance explained by PSTM capacity was 5.2%, which is small but statistically significant. PSTM measured by the forward digit span test significantly predicted L2 reading skills (β = .25, *p* =.001) as measured by the VTR, whereas the pseudoword repetition test did not (β = -.06, *p* = .405).

The results of the standard multiple regression analysis to investigate the amount of variance of L2 grammar knowledge explained by PSTM capacity demonstrate that L2 grammar knowledge as indicated by VTR-Grammar was significantly predicted by the independent variables, F (2, 203) = 3.21, p = .043, with R^2 at .031. The variance explained by PSTM capacity was 3.1%, which is small but statistically significant. The forward digit span test significantly explained the variance in L2 grammar knowledge (β = .19, *p* = .012), whereas the pseudoword repetition test did not (β = -.10, *p* = .210).

Lastly, another standard multiple regression analysis was performed to investigate the amount of variance of L2 text comprehension skills explained by the PSTM. The results showed that L2 text comprehension as indicated by VTR-Comprehension was significantly predicted by the independent variables, F (2, 203) = 6.46, p =.002, with R^2 at .060, and that the variance in text comprehension skills significantly explained by PSTM capacity was 6.0%, which is small but statistically significant. As with other models, only the forward digit span test significantly explained the variance in text comprehension skills (β = .27, *p* < .001), whereas the pseudoword repetition test did not (β = -.10, *p* = .203).

Table 2

Standard Multiple Regression Results for PSTM Predicting the Variables of L2 Reading Sub-Skills

	VTR-Total	VTR-Vocabulary	VTR-Grammar	VTR-Comprehension
β for Forward Digit Span Test	.26**	.25**	.19*	.27***
β for Pseudoword Repetition Test	-.10	-.06	-.10	-.10
R^2	.056	.052	.031	.060
F for change in R^2	6.05**	5.54**	3.21*	6.46**

Note. PSTM = phonological short-term memory; VTR = VELC Test Reading section.

$^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Discussion & Conclusion

In reading, only visual language information is presented. However, the subvocal rehearsal system comprising PSTM, where participants subvocalize the items to be memorized, "serve[s] the function of registering visual information within the store" (Baddeley, 2003, p. 191). In reading L2 text, readers use their phonological knowledge to rehearse the presented language in their mind, holding and processing the meaning of the text. This could be part of the reason for the significant relationship between PSTM and L2 reading skill demonstrated in this study. Despite this theoretical implication, some previous studies discussed above indicated that the effects of PSTM on L2 reading proficiency were not significant. However, this current study shows a significant contribution of PSTM to L2 reading proficiency, although the effect size is not large.

L2 vocabulary and grammatical knowledge were also examined as subskills of L2 reading. The results demonstrated that PSTM influences L2 receptive vocabulary and grammatical knowledge, which supports the results of previous studies (e.g., French & O'Brien, 2008; Martin & Ellis, 2012; Masoura & Gathercole, 1999), and the effect of PSTM on grammatical knowledge was found to be smaller than that on vocabulary knowledge. Completing a sentence using grammatical knowledge, which is a task in VTR-Grammar, requires higher-level cognitive processing than retaining verbal information, which requires PSTM. This might be the reason for the reduced influence of PSTM on L2 grammatical knowledge.

The results show that the largest effect size was for the effects of PSTM on text comprehension, as measured by VTR-Comprehension. To complete the tasks in the VTR-Comprehension part, test-takers need to hold larger amounts of verbal information to memorize than in the VTR-Vocabulary and VTR-Grammar parts, because the number of words in each item is greater than in those of the other tasks. The ability measured in the VTR-Comprehension part was passage comprehension assessed through a context-dependent sentence completion task. Owing to the large amount of information examinees needed to keep in memory, it is plausible that the effect of PSTM was more strongly related to the outcome of this part of the test than those of the other parts.

Although PSTM capacity as indicated by the forward digit span test had a significant influence on L2 reading and related subskills, PSTM as indicated by the pseudoword repetition test did not. One reason for this result might be related to the issues with mishearing the sounds (i.e., *mora*) of the pseudoword. Although test items were constructed with Japanese *mora*, some participants may have misheard the pseudowords; for example, some participants reproduced the items *te-ni-ho* as *te-ni-o*, whereas this type of error was not observed in the forward digit span test. This indicates that the pseudoword repetition test may have involved aural sensitivity in addition to holding speech information, which might have influenced the results.

Although this study demonstrated the effects of PSTM on intermediate EFL learners, future studies would benefit from examining the effects of PSTM capacity on L2 reading skills with a wider range (low to advanced) of proficiency, which would allow for the analysis of proficiency groups at different levels. Future studies might also consider controlling for other factors this study did not control for, such as the amount of exposure to the target language, which could impact L2 reading proficiency. In addition, the pseudoword repetition test did not contribute to explaining variance in L2 reading proficiency scores. As discussed above, it is assumed that the involvement of other factors such as aural sensitivity with the pseudoword repetition test attributed to this non-significant contribution. Therefore, future studies might consider using a different PSTM measure, such as the serial recognition test used by Hummel and French (2016). Furthermore, Swanson et al. (2011) demonstrated that visuospatial memory was a significant predictor of the L2 reading skills of Hispanic children in the United States (whose L1 was Spanish). As first language orthographic features affect learning another language writing system (Akamatsu, 1999; Chikamatsu, 1996), future research could thus focus on investigating the effects of visual memory capacity on the L2 reading proficiency of Japanese EFL learners, who have a different L1 orthographic system than English and Spanish. Such research may provide further insights into the understanding of individual differences in mastering L2 reading skills.

This study attempted to demonstrate the contribution of PSTM capacity as an aptitude factor in L2 reading proficiency. This study contributes to the literature by showing the significant effects of PSTM on L1 Japanese-speaking, intermediate-proficiency, EFL learners, which is the level of most university-level learners (ETS, 2020; MEXT, 2018). The results of this study, then, can be used by language teachers to understand the role of PSTM and the development of L2 reading, and thereby, to design reading tasks that cater to the PSTM differences among their students. One way to do this is to control the amount of text presented to a learner, which in turn may lead to better processing of the text by the learner and, by extension, improved efficacy in teaching.

Akiko Kondo is an Associate Professor at the Hyogo University of Teacher Education. Her interests include short-term working memory in Second Language Acquisition (SLA).

References

Akamatsu, N. (1999). The effects of first language orthographic features on word recognition processing in English as a second language. *Reading and Writing*, *11*(4), 381–403. https://doi.org/10.1023/A:1008053520326

Atkins, P. W. B., & Baddeley, A. D. (1998). Working memory and distributed vocabulary learning. *Applied Psycholinguistics, 19*(4), 537–552. https://doi.org/10.1017/S0142716400010353

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417–423. https://doi.org/10.1016/S1364-6613(00)01538-2

Baddeley, A. D. (2003). Working memory and language: An overview. *Journal of Communication Disorders, 36*(3), 189–208. https://doi.org/10.1016/S0021-9924(03)00019-4

Baddeley, A. D. (2010). Working memory. *Current Biology, 20*(4), R136–R140. https://doi.org/10.1016/j.cub.2009.12.014

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–89). Academic Press. https://doi.org/10.1016/S0079-7421(08)60452-1

Cheung, H. (1996). Nonword span as a unique predictor of second-language vocabulary learning. *Developmental Psychology, 32*(5), 867–873. https://doi.org/10.1037/0012-1649.32.5.867

Chikamatsu, N. (1996). The effects of L1 orthography on L2 word recognition: A study of American and Chinese learners of Japanese. *Studies in Second Language Acquisition, 18*(4), 403–432. https://doi.org/10.1017/S0272263100015369

Educational Testing Service (ETS). (2020). 2019 report on test takers worldwide: TOEIC Listening & Reading Test. https://www.iibc-global.org/library/default/toeic/official_data/pdf/Worldwide2019.pdf

French, L. M. (2006). *Phonological working memory and second language acquisition: A developmental study of Francophone children learning English in Quebec*. Edwin Mellen Press.

French, L. M., & O'Brien, I. (2008). Phonological memory and children's second language grammar learning. *Applied Psycholinguistics, 29*(3), 463–487. https://doi.org/10.1017/S0142716408080211

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology, 40*(2), 177–190. https://doi.org/10.1037/0012-1649.40.2.177

Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition, 14*(1), 25–38. https://doi.org/10.1017/S0272263100010457

Hummel, K. M. (2009). Aptitude, phonological memory, and second language proficiency in nonnovice adult learners. *Applied Psycholinguistics, 30*(2), 225–249. https://doi.org/10.1017/S0142716409090109

Hummel, K. M., & French, L. M. (2010). Phonological memory and implications for the second language classroom. *The Canadian Modern Language Review, 66*(3), 371–391. https://doi.org/10.3138/cmlr.66.3.371

Hummel, K. M., & French, L. M. (2016). Phonological memory and aptitude components: Contributions to second language proficiency. *Learning and Individual Differences, 51*, 249–255. https://doi.org/10.1016/j.lindif.2016.08.016

JACET Basic Word Revision Committee. (2003). *JACET list of 8000 basic words*. Japan Association of College English Teachers.

Kormos, J., & Sáfár, A. (2008). Phonological short-term memory, working memory, and foreign language performance in intensive language learning. *Bilingualism: Language and Cognition, 11*(2), 261–271. https://doi.org/10.1017/S1366728908003416

Kumazawa, T. (2015). *VELC Test no gaiyou to yokuaru shitsumon: Reading Section Part 3 no sakumon ito to koumoku tokusei*. [Frequent questions from the VELC Test overview: Aims and features of items in Reading Section Part 3]. https://www.velctest.org/group/2015/doc_kumazawa.pdf

Kumazawa, T., Shizuka, T., Mochizuki, M., & Mizumoto, A. (2016). Validity argument for the VELC Test® score interpretations and uses. *Language Testing in Asia*, *6*(2), 1–18. https://doi.org/10.1186/s40468-015-0023-3

Martin, K. I., & Ellis, N. C. (2012). The roles of phonological short-term memory and working memory in L2 grammar and vocabulary learning. *Studies in Second Language Acquisition, 34*(3), 379–413. https://doi.org/10.1017/S0272263112000125

Masoura, E. V., & Gathercole, S. E. (1999). Phonological short-term memory and foreign language learning. *International Journal of Psychology, 34*(5/6), 383–388. https://doi.org/10.1080/002075999399738

Ministry of Education, Culture, Sports, Science, and Technology (MEXT). (2018). Heisei 29 nendo eigoryoku chousa kekka (koukou 3 nensei) [2017 survey results of English proficiency (12th grade students)]. https://www.mext.go.jp/a_menu/kokusai/gaikokugo/__icsFiles/afieldfile/2018/04/06/1403470_03_1.pdf

Nakanishi, H. (2011). The relation of L2 working memory capacity of Japanese EFL learners with the resolution of lexical ambiguity. *Journal of the Tohoku English Language Education Society, 31,* 93–103.

O'Brien, I., Segalowitz, N., Collentine, J., & Freed, B. (2006). Phonological memory and lexical, narrative, and grammatical skills in second language oral production by adult learners. *Applied Psycholinguistics, 27*(3), 377–402. https://doi.org/10.1017/S0142716406060322

Watanabe, A. (2018). Koukou no eigo kyoukasho mo bunpou henchou kara 4ginou juushi he [The focus of English textbooks will change from grammar to integrating four skills]. https://benesse.jp/eigo/201803/20180312-1.html

Service, E. (1992). Phonology, working memory, and foreign-language learning. *The Quarterly Journal of Experimental Psychology, 45A*(1), 21–50. https://doi.org/10.1080/14640749208401314

Service, E., & Kohonen, V. (1995). Is the relation between phonological memory and foreign language learning accounted for by vocabulary acquisition? *Applied Psycholinguistics, 16*(2), 155–172. https://doi.org/10.1017/S0142716400007062

Shizuka, T., & Mochizuki, M. (2014). VELC Test for testing competency: Verification of reliability and validity. http://www.velctest.org/contact/VelcTest-for-TestingCompetency.pdf

Swanson, H. L., Orosco, M. J., Lussier, C. M., Gerber, M. M., & Guzman-Orth, D. A. (2011). The influence of working memory and phonological processing on English language learner children's bilingual reading and language acquisition. *Journal of Educational Psychology, 103*(4), 838–856. https://doi.org/10.1037/a0024578

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Pearson Education.

van den Noort, M. W. M. L., Bosch, P., & Hugdahl, K. (2006). Foreign language proficiency and working memory capacity. *European Psychologist, 11*(4), 289–296. https://doi.org/10.1027/1016-9040.11.4.289