The Contribution of Lexical Processing Ability to Efficient Functioning of Working Memory in Japanese EFL Learners

SUGIURA, Kaori
Shizuoka University of Art and Culture

NAKANISHI, Hiroshi
Tohoku Gakuin University

要旨

人が文字を見て理解するまでの過程には、文字認知・音韻符号化・単語の意味理解・統語処理（低次処理）・意味処理・文脈処理・スキーマ処理（高次処理）といった様々な処理過程が含まれる。それぞれの処理はワーキングメモリ（WM）上で行われ、各処理段階でWMが消費されると考えられている。そのWM容量には厳しい容量制限があるため、低次での言語処理をいかに効率よく行い、高次処理段階にWM資源を分配できるかが、言語理解の鍵になる。特に、言語処理が自動化していない日本人英語学習者にとっては、WMの効率性が言語理解に果たす役割はＬ１よりも大きいと考えられる。本研究の目的は、低次処理である語彙処理に焦点を当て、語彙知識量と語彙処理の効率性がWMの運用効率に果たす役割について検討することにある。本実験では30人の日本人英語学習者を対象に、語彙知識量を測るVocabulary Size Test（VST）(Nation & Begler, 2007）と語彙処理能力を測るために開発された語彙処理テスト（CELPテスト）(Kadota, 2010)、WM容量を測定するためのリーディングスパンテスト（RST）(Nakanishi, 2005）を行い、各テスト成績間における相関関係を調査した。その結果、ある一定量の語彙知識を持った学習者においては、語彙処理能力とWM容量の間に相関傾向がみられたが、一定量以下の語彙知識を持つ学習者においては、語彙知識量とWM容量の間に有意な相関がみられた。

Key words：lexical access, vocabulary size, working memory, second language learners
キーワード：語彙アクセス、語彙サイズ、ワーキングメモリ、第二言語学習者

1. Introduction

In language comprehension, visual or aural information is processed in working memory (WM) by retrieving appropriate information from our knowledge in long-term memory (i.e., lexical, syntactic, and semantic knowledge) (Baddeley, 1986). The processed information is maintained for a short time and integrated with a series of inputs to WM. In this way, WM plays an important role in language comprehension, supporting both the processing and storage functions (Just & Carpenter, 1992).

WM is thought to be a limited cognitive resource; both processing and storage of relevant information are assumed to draw on the same resource supply. When task demands exceed the resource capacity, a tradeoff relationship exists between the processing and storage functions. This tradeoff results in major negative consequences for task performance, leading to slower processing and the loss of information (Just & Carpenter, 1992). Therefore, retrieving appropriate information from long-term memory using as few WM resources as possible is a key factor in the efficient functioning of WM because this reserves a large portion of WM resources for other processing and for storage of information.

The automaticity of lexical access is important to the efficient use of WM, and to language comprehension. Although a number of researchers have focused mainly on learners' amount of vocabulary knowledge, in everyday communicative situations, the efficient retrieval of vocabulary knowledge is more crucial. Based on this concept, Kadota (2010) recently developed the Computer-based English Lexical Processing Test (CELP Test)[1] to measure how accurately learners processed vocabulary, as well as how quickly students accessed their mental lexicons.

The purpose of this study is to examine the way in which efficient lexical access contributes to the efficient use of WM by comparing with the relationship between vocabulary size and WM efficiency.

2. Literature Reviews

2.1 Working memory capacity and language comprehension

Language comprehension consists of both lower and higher-level processing. Lower-level processing includes word recognition, phonological encoding, and lexical access; while higher-level processing covers meaning, sentence context, schema, and discourse processing (Kadota, 2007). Automaticity in lower-level processing conserves sufficient WM capacity for higher-level processing.

On a first language reading study, Samuels (1994) shows how during reading, beginning and fluent readers allocate their attentional resources to both decoding and comprehension processes (see Figure 1). The figure indicates that for beginners, the decoding process is extremely resource demanding. Hence, they might be unable to allocate sufficient resources to the comprehension process. On the other hand, fluent readers perform the decoding process with complete automaticity, leaving their attentional resources available for the comprehension process.

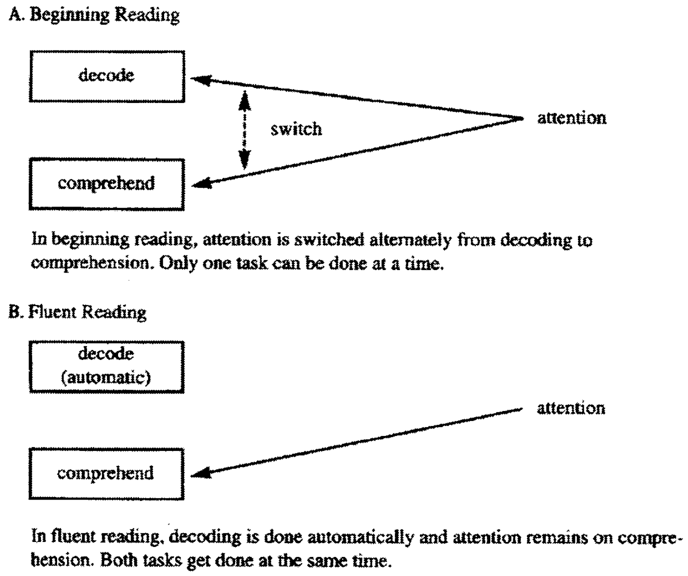

Figure 1. The relationship between attention and reading (Samuels, 1994)

A number of L1 reading studies have indicated that the efficiency of lower-level processing skill is strongly related to successful reading performance supported by the balanced availability of WM resources (e.g., Daneman & Carpenter,1980).

On the other hand, there have been few second language (L2) studies that have examined how lower-level processing contributes to L2 WM capacity. Kato (2003) investigated the relationship for English as foreign learners between their low-level processing skills and WM capacity. In an orthographic processing task, Kato (2003) asked participants to examine a pair of non-words (e.g., taidge–dgait) and identify the one that most resembled a real English word. The results revealed that orthographic processing skills highly correlated not only with reading comprehension but also with L2 WM capacity. The results suggested that the effectiveness of the orthographic processing skills could help conserve a larger amount of WM resources for higher-level processing. Given this fact, efficient lexical access—also one of the lower-level processing skills—might afford sufficient resources for higher-level processing.

2.2 Vocabulary tests

It is well known that learners' vocabulary knowledge has been measured based on "breadth" and "depth" (Read, 1993). According to Read (1993), vocabulary breadth refers to the total number of words that learners know, whereas vocabulary depth refers to how much learners know about target words.

Regarding vocabulary tests for measuring vocabulary breadth, Vocabulary Levels Test (VLT) was developed by Nation (1983), and revised by Schmitt, Schmit and Clapham (2001). This test was originally developed for the purpose of checking learners' vocabulary levels (Kadota, 2010). Therefore, Nation and Beglar (2007) created Vocabulary Size Test (VST), which has the function of measuring learners' knowledge of vocabulary breadth. As for the tests targeting Japanese English language learners, a number of vocabulary tests have been developed (e.g., Aizawa, 2007 for J8VST).

There have been several tests developed for measuring vocabulary depth. Some examples of these tests include, Lex 30 (Meara & Fitzapatric, 2000), which is based on word association tasks, and V_links (Meara & Wolter, 2004), which measures the density of vocabulary networks. Similar to V_links, Mochizuki (2010) have developed the Lexical Organization Test (LOT), which targets Japanese learners.

The vocabulary tests described above measure learners' vocabulary knowledge in terms of "breath" and "depth," and have the purpose of measuring learners' declarative knowledge of vocabulary. By considering the necessity of vocabulary use in communication, it is crucial to develop vocabulary tests that can assess learners' ability to put vocabulary knowledge into actual use—that is, procedural knowledge of vocabulary. Thus, the CELP test was developed by Kadota (2010), which measures efficiency and automaticity of vocabulary processing, such as learners' access speed to target words.

This study investigates the relationship between vocabulary skills and efficient use of WM capacity. To measure the two types of vocabulary skills, the VST (Nation & Beglar, 2007) and the CELP test (Kadota, 2010) were employed. As previously mentioned, the VST evaluates the size of learners' vocabularies, whereas the CELP measures the efficiency and accuracy of lexical access. As for the measurement of efficient use of WM capacity, the Reading Span Test (RST) was employed. This study will reveal the contribution of automaticity in lexical access to the efficiency of WM.

3. Methods

3.1 Participants

The participants for this experiment were 30 Japanese university students—21 females and 9 males.

3.2 Procedure

All participants completed the following tasks: (1) the RST, (2) the VST, and (3) the CELP Test. All the tests were administered for one student participant at a time. The order of the tasks was counterbalanced in this experiment. The entire experiment took approximately 1 hour for each participant.

3.2.1 Reading span test

The RST, which was originally developed by Daneman and Carpenter (1980), has been used to measure a participant's verbal WM capacity. This experiment used a revised version of the RST for Japanese EFL learners (Nakanishi, 2005).

The RST was implemented on a Windows computer, using a psychological experiment software package called SuperLab Pro (Cedrus Corporation). The procedure was as follows:

- A fixation marker was presented for 1 second on a computer monitor, and then the marker was replaced with a sentence.
- The participant was asked to read the sentence aloud and then press the space key immediately after finishing the sentence. The participant was also asked to remember the final word in the sentence.
- After the participant pressed the space key, the Japanese equivalent of the English sentence appeared on the monitor.
- The participant was asked to judge whether the Japanese equivalent was accurate, and to press the (B) key if accurate or the (N) key if not.
- The reaction time between the emergence of the question and when the participant pressed the space key was recorded.
- The presentation and response cycle then started again, with a fixation mark followed by the next sentence.
- This procedure repeated until the participant received an onscreen instruction

indicating the end of the computerized session.

- Then the participant was asked to write on an answer sheet the sentence-ending words that had been presented. For example, under the three-sentence condition, the participants read three sentences and tried to remember the sentence-ending words, then encountered the instruction showing the end of the session. They were then asked to write the three final words on the answer sheet.

The sets of sentences consisted of two to five sentences, 14 sentences in all, which were presented in order, smallest set to largest set. The sentences ranged from nine to 13 words long. There were three sessions, which added up to 42 recall words. The sentences were selected from those used in the L2 studies by Osaka and Osaka (1992) and Harrington and Sawyer (1992). We made minor modifications so that among the sets, the familiarity of the words (Yokokawa, 2006) was statistically the same, on average (F = 1.1445, *n.s.*).

3.2.2 Vocabulary size test

In this study, we employed the VST that Nation and Beglar (2007) developed to measure L2 learners' breadth of vocabulary knowledge. It can be said that this test is the one of the most well known vocabulary size tests in the world.

The full VST consists of 14 levels, ranging from a 1,000-word level to a 14,000-word level, and each level contains 10 items (Beglar, 2010; Nation & Beglar, 2007). It is comprised of 14,000 of the highest frequency vocabulary words from the spoken corpus of the British National Corpus (Leech, Grayson, & Wilson, 2001).

For the current study, we used the words in the first six levels (1,000-word level to 6,000-word level), a total of 60 words. On the VST, each item was presented in a decontextualized sentence, with four possible definitions, including one correct answer and three distractors. There was a time limit of 30 minutes. Figure 2 shows an example from the VST:

Third 1000

1.	soldier: He is a **soldier**.	
	a.	person in a business
	b.	student
	c.	person who uses metal
	d.	person in the army

Figure 2. An example from the VST, Third 1000 (Adapted from Nation & Beglar, 2007)

3.2.3 The Computer-based English lexical processing test

The CELP Test was developed by Kadota (2010) to measure the speed and accuracy of L2 learners' access to their mental lexicons. The test was implemented on a Windows computer using Visual Basic 2005. The procedure was as follows:

- After the fixation marker was presented for 2,000 ms on a computer monitor, the marker was replaced with a prime English word for 300 ms.
- The prime word was replaced by the target English word for 100 ms.
- The target word was then replaced by a blank screen.
- The participants were asked to judge whether the two words were semantically similar and then to press the corresponding letter key: (B) for similar, (N) if not.
- The system recorded the participants' reaction time from the emergence of the target word to the press of the response key.
- Then the next cycle began with the fixation mark and new prime word. This procedure was repeated until the participant received an onscreen instruction indicating the end of the session.
- Prior to the experimental phase, participants completed 12 practice trials and received feedback for first eight of these practice trials.

The CELP presents 100 word pairs, including 50 semantically similar word pairs and 50 semantically dissimilar word pairs. Table 1 indicates examples of the paired items in the CELP test. The word pairs were selected from a list of 3,000 familiar words for Japanese EFL learners (Yokokawa, 2006). The word familiarity refers to the degree of how much the leaners feel that they hear or see the targeted words.

For the CELP test, higher-familiarity words were selected as prime words rather than target words. The words used for the CELP test are considered to be in accordance with those of the 1,000 and the 2,000 levels in the VST (Kadota, 2010). Thus, it can be assumed that the words the participants encountered in the CELP were easier than those in VST.

Table 1

Examples of the paired items in the CELP Test

Prime (familiarity)	Target (familiarity)
get (6.60)	acquire (4.16)
carry (6.00)	convey (3.23)
true (6.33)	genuine (2.93)

Note. ( ) = familiarity rated on a 7-point scale (1 = unknown, 7 = very familiar; Yokokawa, 2006).

4. Results

4.1 Data descriptions of tests (RST, VST, and CELP) for all participants

First, Table 2 shows reading span scores. The Recall score is the total number of sentence-ending words remembered, the e-f score represents the number of recalled words when the participants correctly processed the presented sentences, and Accuracy shows the number of sentences processed correctly.

Table 2

Descriptive statistics of RST data for all participants

	Recall	e-f score	Accuracy
Avg.	27.30	20.27	31.57
Min.	16.00	7.00	20.00
Max.	39.00	33.00	39.00
S.D.	4.36	6.24	5.34

Note. N = 30. Mark range: 0–42 (Recall, e-f score, Accuracy).

Table 3 presents descriptive statistics for the VST scores, including each score from the 1,000-word level to the 6,000-word level, along with the total score for all participants. The maximum score for each level is 10; therefore, the maximum total score is 60.

Table 3

Descriptive statistics of VST scores for all participants

Level	1000	2000	3000	4000	5000	6000	Total score
Avg.	8.13	5.87	6.07	5.97	4.63	3.57	49.82
Min.	5.00	1.00	3.00	2.00	1.00	1.00	29.85
Max.	10.00	10.00	9.00	10.00	9.00	7.00	72.25
S.D.	1.12	2.06	1.55	1.45	1.87	1.41	9.95

Note. $N = 30$. Mark range: 0–10 (1,000–6,000), 60 (total).

Table 4 provides the average number of errors and the reaction times in the CELP Test for all participants. These scores are converted into a Z-score.

Table 4

Descriptive statistics of CELP for all participants

	Error	Time	Z-score
Avg.	0.14	1.21	50.00
Min.	0.01	0.80	36.67
Max.	0.26	2.17	69.16
S.D.	0.06	0.32	7.82

Note. $N = 30$. Mark range: 0–1 (Error, Time).

4.2 VST,CELP, and RST data by vocabulary-size group

Thirty participants were divided into two groups (15 participants for each group), according to their VST total scores. Those with a larger vocabulary were in one group, while those with a smaller vocabulary were in another. Table 5 provides the total VST data for both the larger and smaller groups. The average VST total score of the larger vocabulary group was 56.61 out of 60.00 points, whereas that of the smaller group was 43.04 points. A *t*-test showed that the larger vocabulary group achieved significantly higher scores than the smaller vocabulary group: $t(28) = 6.186, p < .05$.

Table 5

VST data by vocabulary-size group

	Large	Small
	VST Total	VST Total
Avg.	56.61	43.04
Min.	43.98	29.85
Max.	72.25	55.29
S.D.	7.14	7.42

Table 6 shows the CELP data for individuals with larger and smaller vocabulary sizes. A *t*-test indicated that there was a significant difference between the large and small vocabulary groups: $t(28) = -4.174$, $p < .01$.

Table 6

CELP data by vocabulary-size group

	Large	Small
	CELP Z-score	CELP Z-score
Avg.	45.16	54.84
Min.	36.67	47.27
Max.	60.12	69.16
S.D.	5.38	6.81

Note. A Z-score was calculated based on the average number of errors and the reaction times in the CELP Test for all participants.

Table 7 shows the RST data for individuals with larger and smaller vocabulary sizes. A significant difference exists between the large and small vocabulary-size groups in Recall and e-f scores (Recall: $t(28) = 2.241$, $p < .05$; e-f score : $t(28) = 2.351$, $p < .05$).

Table 7

RST data by vocabulary-size group

	Large		Small	
	Recall	e-f score	Recall	e-f score
Avg.	29.00	22.80	25.60	17.73
Min.	20.00	11.00	16.00	7.00
Max.	39.00	33.00	30.00	26.00
S.D.	4.56	5.62	3.38	5.78

4.3 Correlational analyses

Table 8 shows the Pearson Product Moment correlation coefficients calculated among the RST, VST, and CELP scores. In the RST scores, there was a high correlation between Recall and e-f scores (r = .85, p < .01). In the vocabulary tests, there was a relatively high correlation between the VST and CELP scores (r = –.66, p < .01). As for correlations between RST and vocabulary performances, RST scores produced relatively high correlations with VST and CELP scores regardless of scoring methods (Recall and VST: r = .53, p = < .01; Recall and CELP: r = –.48, p = < .01; e-f score and VST: r = .60, p = < .01; e-f score and CELP: r = –.44, p < .05).

Table 8

Correlational analysis for all participants

	RST		Vocabulary	
	Recall	e-f score	VST	CELP
Recall	—	**0.85	**0.53	**–0.48
e-f score		—	**0.60	*–0.44
VST			—	**–0.66
CELP				—

Note. N = 30. *p < .05. **p < .01.

The correlation analysis was conducted for the larger-vocabulary group. As shown in Table 9, CELP scores tended to produce a significant correlation with Recall and e-f scores (r = .48, p < .01 and r = –.44, p < .05, respectively), whereas VST did not produce any significant correlation with Recall and e-f scores (r = .32, *n.s.* and r = .42, *n.s.*, respectively).

Table 9

Correlational analysis for larger-vocabulary group

	RST		Vocabulary	
	Recall	e-f score	VST	CELP
Recall	—	**0.86	0.32	†-0.46
e-f score	—	—	0.42	†-0.50
VST				-0.42
CELP				—

Note. $N = 15$. †$p < .10$. *$p < .05$. **$p < .01$.

Table 10 shows the correlation analysis conducted for the smaller-vocabulary group. The results indicate a significant correlation between VST and RST scores (Recall: $r = .55$, $p < .05$; e-f score: $r = .62$, $p < .05$), whereas there was no significant correlation between CELP and RST scores (Recall: $r = -.21$, *n.s.*; e-f score: $r = -.08$, *n.s.*).

Table 10

Correlational analysis for smaller-vocabulary group

	RST		Vocabulary	
	Recall	e-f score	VST	CELP
Recall	—	**0.81	*0.55	-0.21
e-f score		—	*0.62	-0.08
VST			—	-0.36
CELP				—

Note. $N = 15$. †$p < .10$.

5. Discussion and Conclusion

The main results of the present experiment are as follows:

- For participants with large vocabularies, RST scores tended to correlate significantly with CELP scores.
- For participants with small vocabularies, RST scores correlated significantly with VST scores.
- RST scores were higher for the group of participants with larger vocabularies, regardless of scoring methods.

The correlation results imply that for the group with a larger-vocabulary-size, the degree of automatization of lexical processing determines the efficient functioning of WM, whereas for the group with a smaller-vocabulary-size, their vocabulary size is a key factor that helps function WM efficiently.

As many previous L1 studies have pointed out, efficient functioning of WM is a key factor in language comprehension. For Japanese EFL learners, whose language processing in English is less automatized than in L1 processing, L2 processing will consume more WM resources than L1 processing and will rapidly deplete those resources. Therefore, the efficient functioning of WM could be a stronger determining factor for language comprehension in L2 than in L1 processing (Geva & Ryan, 1993; Miyake & Friedman, 1998). In particular, automatization of low-level processing, such as lexical access, is important because it reserves WM resources for higher-level processing.

The present study provides good pedagogical indicators for improving efficiency in the use of WM resources in terms of vocabulary instruction. For students with larger vocabularies (i.e., the average VST score: 56.61 out of 60.00 points), language teachers should encourage students to retrieve their knowledge efficiently and unconsciously, whereas for students with smaller vocabularies (i.e., the average VST score: 43.04 out of 60.00 points), teachers should first work on increasing the students' vocabularies. In this way, language teachers should be flexible in changing their instructions depending on students' vocabulary sizes, and the educational focus should shift gradually from developing vocabulary size to fluent use of vocabulary as vocabulary size increases.

Notes

1 CELP was developed as part of the Grant-in-Aid for Scientific Research (C) "The interface between lexical and sentence processing in L2: An empirical study of Japanese EFL learners" (PI: Shuhei Kadota, No. 19520532,). The authors participated in the research project as collaborators.

Acknowledgements

Our thanks go to our students who helped us with our research by giving valuable data. We are grateful to the anonymous reviewers for their valuable comments and suggestions.

This study was partially supported by a JSPS Grant-in-Aid for Scientific Research (C)

"The interface between lexical and sentence processing in L2: An empirical study of Japanese EFL learners" (PI: Shuhei Kadota, No. 19520532) and Grant-in-Aid for Research Activity Start-up "The relation of Working Memory capacity of Japanese EFL learners with the processing of syntactically ambiguous sentences" (PI: Hiroshi Nakanishi, No. 22820052).

References

Aizawa, K. (2007). *Construction of a Vocabulary List for Japanese Learners of English and Development of a System for Analyzing Educational Materials Based on Large-scale Corpora.* (supported by a JSPS Grant-in-Aid for Scientific Research (B)) (PI: Aizawa Kazumi, No. 16320076). Retrieved from www.j-varg.sakura.ne.jp/download/.../kaken_(Aizawa)_H16-18.pdf

Baddeley, A. D. (1986). *Working memory.* New York: Oxford University Press.

Beglar, D. (2010). A rasch-based validation of the vocabulary size test. *Languag Testing, 27,* 101-118.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450-466.

Geva, E., & Ryan, E. B. (1993). Linguistic and cognitive correlates of academic skills in first and second languages. *Language Learning, 43,* 5-42.

Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition, 14,* 25-38.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99,* 122-149.

Kadota, S. (2007). *Science of shadowing and reading aloud.* [*Shadowing to ondoku no kagaku*]. Tokyo: Cosmopier.

Kadota, S. (Ed.). (2010). *The interface between lexical and sentence processing in L2: An empirical study of Japanese EFL learners. KAKEN final report.* (supported by a Grant-in-Aid for Scientific Research (C)). (PI:Kadota Shuhei, No. 19520532).

Kato, S. (2003). Examining relationships between phonological/orthographic processing efficiency, working memory, and overall reading performance: Implications for developmental changes in L2 reading proficiency. *JACET Bulletin, 37,* 31-48.

Leech, G., Grayson, P., & Wilson, A. (2001). *Word Frequencies in Written and Spoken*

English: based on the British National Corpus. London: Longman.

Meara, P. M., & Fitzpatrick, T. (2000). Lex 30: an improved method of assessing productive vocabulary in an L2. *System, 28,* (1). 19-30.

Meara, P. M., & Wolter, B. (2004). V_Links: Beyond Vocabulary Depth. *Angles on the English Speaking World, 4,* 85-97.

Miyake, A., & Friedman, N. P. (1998). Individual differences in second language proficiency: working memory as language aptitude. In A. Healy & L. Bourne (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp.339-365). London: Lawrence Erlbaum Associates.

Mochizuki, M. (2010). *Development of a Vocabulary Test Battery Estimating English Skills and Proficiency: Integrating Vocabulary Size, Organization, and Access Speed.* (supported by a JSPS Grant-in-Aid for Scientific Research (B)). (PI: Mochizuki Masamichi, No. 19320084). Retrieved from http://www.j-varg.sakura.ne.jp/download/materials/Kaken(Mochizuki)_H19-21.pdf

Nakanishi, H. (2005). How memory spans for reading, listening and spatial thinking are related with reading and listening comprehension for Japanese EFL learners. *Language Communication and Culture, 3,* 123-137.

Nation, P. (1983). Learning vocabulary. *New Zealand Language Teacher, 9,* 1, 10-11.

Nation, P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher, 31,* 7, July, 9-12.

Osaka, M., & Osaka, N. (1992). Language-independent working memory as measured by Japanese and English reading span tests. *Bulletin of the Psychonomic Society, 30,* 287-289.

Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. *Language Testing, 10,* 3, 355-371.

Samuels, S. J. (1994). Toward a theory of automatic information processing in reading, revised. In R. B. Ruddell, M. R. Ruddell & H. Singer (Eds.), *Theoretical models and processes of reading, 4,* (pp.816-837). Newark, DE: International Reading Association.

Schmitt, N., Schmitt,D., & Clapham, C. (2001). Developing and exploring the behavior of two new versions of the Vocabulary Levels Test. *Language Testing, 18,* 1, 55-88.

Yokokawa, H. (Ed.). (2006). *Familiarity of English words for Japanese EFL learners,* *Script version – Second language database for education and research.* [*Nihonjin eigo gakushusha no eitango shinmitsudo moji hen – kyoiku kenkyu no tameno dai-nigengo datebase*]. Tokyo: Kuroshio Publisher.