# *Effects of Extensive Reading on Reading Fluency: Evidence From Word Recognition and Syntactic Processing*

**NORO, Tadashi , ROHE, Gregory Lewis, OKADA, Sayuri, HASEBE, Ako**

## Introduction

In Japanese high school classrooms surprisingly little English reading is done across the curriculum (Kanatani, 2008; Tono, 2015). Learners are required to read rather difficult textbooks slowly, stopping constantly to look up unknown words in the dictionary. Such "slow readers seldom develop much interest in what they read, let alone pleasure. Since they do not enjoy it, they read as little as possible. Deprived of practice, they continue to find it hard to understand what they read, so their reading rate does not increase. They remain slow readers" (Nuttall, 2005, p.127). Many of our learners are trapped in this vicious circle. On the other hand, learners in the virtuous circle read independently for enjoyment, and read a lot of texts faster, with more processing experience, becoming more efficient in comprehending texts better. Thus, speed, enjoyment, quantity and comprehension of reading are closely linked, and provide the key to raise learners out of the vicious circle and into the virtuous circle. These factors are attained by the amount of reading practice, or *extensive reading*, which will accelerate reading fluency development (Nuttall, 2005, p. 128).

"Fluency development is often neglected in courses, partly because teachers and learners feel that they should always be learning something new" (Nation, 2009, p.2). Actually, fluency activities are seldom practiced in most L2 curricula and minimized in most L2 reading textbooks, mainly because the goal for a reading class is the development of language skills, vocabulary, grammar, translation ability, or study skills, and probably because "the development of accurate, even if very slow, comprehension abilities is the real goal of many L2 reading curricula, with the assumption that a good comprehender will eventually become a fluent reader of extended texts on his or her own if he or she is so inclined or needs this ability" (Grabe, 2009, p.312). Actually, however, "it is necessary to note that fluent reading is what most good readers do most of the time when they read almost anything, especially in their L1" (Grabe, 2010, p.72). Effective L1 reading comprehension generally assumes reading fluency. A fluent L1 reader is "a person reading at a reasonable reading rate, between 250-300 WPM, using very efficient and fast word recognition skills, and combining information from various sources while reading under fairly intense time constraints" (Grabe, 2010, p.72). The basic concept of L1 reading fluency is relevant to skilled L2 readers. Where L2 learners read very easy books with virtually no unknown vocabulary and grammatical items, they are exposed to the same expressions repeatedly, and thus can make the best use of what is already known; extensive reading provides the conditions for fluency development. The readers develop automatic word-recognition skills, expand the breadth and depth of vocabulary knowledge, have autonomous control over syntactic-parsing skills, and read for additional learning. From such efficient processes, it is expected that "when fluency practice is incorporated as a key component of a reading curriculum, students make important gains in their reading comprehension" (Grabe & Stoller, 2011, p. 150).

How extensive reading can contribute to fluency development in reading is best explained from the cognitive perspective. The perspective mainly suggests that to comprehend the meaning of a given text efficiently, the processes of lower-level components of reading, such as word recognition and syntactic processing, must be automatized (Samuels, 2006; Perfetty, 1985). In order to clarify the development of lower-level processes through extensive reading, the present study will first investigate the speed, as well as the accuracy, of word recognition. Second, it will examine the speed, as well as the accuracy of syntactic processing. Third, it will look into relationships of word recognition and syntactic processing with reading comprehension.

## Literature Review

### Defining Reading Fluency

What does reading fluency mean? As Kame'enui and Simmons (2001, p. 204) say, "reading fluency does not enjoy definitional, theoretical, or instructional consensus in the research literature." Defining reading fluency is a complex issue. Concepts of reading fluency differ among different researchers. Breznitz (2006) mentions three viewpoints on the definition of reading fluency: (1) fluency as an outcome of the quality of oral reading skills (pp. 4-5), (2) fluency as an outcome of the development in each component (pp. 5-7), and (3) fluency as an outcome of the effectiveness of various biological and cognitive systems (p. 8). By 'the effectiveness of biological and cognitive systems', she means "fine-timing mechanisms in the brain to synchronize information for comprehension" (cited from Grabe, 2009, p. 300).

Grabe (2009, p. 291) defines fluency in reading as (1) "the ability to read rapidly with ease and accuracy and (2) to read with appropriate expression and phrasing," involving (3) "a long incremental learning process, and (4) text comprehension as an expected outcome of fluent reading." Considering that comprehension is a capacity-limited cognitive process, Grabe (2009) adds automaticity to accuracy and rate as central subprocesses of fluent reading, because automaticity allows readers to attend to the meaning of the text, the textual context and required background knowledge without being slowed down by attentional word-recognition demands (Samuels, 2002; Segalowitz, 2000, National Reading Panel, 2000). In discussion of fluency, automaticity is defined as processing operations that are rapid, relatively resource-free, not subject to interference, unconscious, and hard to oppress (National Reading Panel, 2000), and that are fast, ballistic, load independent, effortless, and unconscious (Segalowitz, 2003, pp. 385-394) . Besides, Grabe (2010, p. 73) claims that lower-level component skills of reading － word recognition, a large recognition vocabulary, skilled grammatical processing, and the formation of basic meaning proposition units for reading comprehension － are implicitly learned gradually over a long period of time. Those skills only "emerge as an outcome of implicit learning through extended periods of exposure and meaningful time on task … Implicit learning gradual, initially very fragile, and strongly based in repetition of form and process through extensive reading over a long period of time" (Ellis, 2005, 2007). This is why the connection between fluency and implicit learning clearly points to the pedagogical importance of extensive reading (Grabe, 2009; Nation, 2009).

Based on the discussion above, three subprocesses in reading fluency － automaticity, accuracy and rate － are common and central (Kuhn & Stahl, 2003). As for automatization, Segalowitz and Segalowitz (1993) distinguish between "speedup" (mean reaction time [RT] and mean standard deviation decrease in an RT task to the same degree) and "automatization" (mean standard deviation decreases more than mean RT). The coefficient of variation [CV], which is the standard deviation divided by the mean RT, decreases in the case of automatization while remaining unchanged in the case of speed up. Whether or not CV is a reliable index of the acquisition of automatic processing skills is controversial. Some empirical studies (e.g Harrington, 2006, Akamatsu, 2008) partially supported their automatization theory, but others (e.g. Hulstijn, Van Gelderen, & Schoonen, 2009) did not. Besides, since the participants in this study are exposed to a considerable amount of comprehensible input for the first time, it is predictable that they can gain processing skills, but not automatic processing skills. Therefore, this study focuses on accuracy and rate. A similar concept of reading fluency is also grasped by other scholars, as being a combination of an accuracy component and a rate component (Carver, 1997) and as the ability to read accurately and rapidly (Kame'enui & Simmons, 2001). As to oral reading skills or prosodic structuring, they are excluded from this research, since the present writers are concerned not with oral reading but extensive reading, though oral reading skills are closely connected with word recognition, and prosodic structuring with grammatical and semantic processing.

## Why Is Fluency Important in Reading?

Current versions of reading models are interactive in that they view reading processes to be cyclical rather than linear, to work independently of one another, and to operate in a parallel manner (Rumelhart, 1977). Stanovich (1980) viewed reading comprehension as an outcome of processes "based on information provided simultaneously from several knowledge sources" (p. 35). In addition to this, to account for individual differences in reading, Stanovich supplemented his model with a compensatory mechanism: the idea that information at one level of processing might influence and compensate for deficiencies at other levels of processing. For the sake of simplicity, based on reading "component skills approaches" (Carr & Levy, 1990), reading processes are divided into two parts: lower-level processes and higher-level processes. In understanding the fluent reading process, it is essential to focus on lower-level processes. These processes include word recognition, syntactic processing and semantic proposition encoding. These processes are carried out as part of working memory, in which cognitive processing and knowledge resources are integrated for comprehension (Baddeley, 1986). So, why is it important to recognize the role played by lower-level processes? Perfetti and Lesgold (1977) claim that there is a capacity limitation in working memory, as employed in the "limited-capacity model," and there is a trade-off of cognitive resources among multiple components of reading processing. This model explains that as one component of reading processes becomes automatic, more cognitive resources are allocated to the other reading processing components. Samuels (2006) pointedly illustrates the role of decoding automaticity in overall comprehension. The illustration indicates that if word recognition becomes automatic, more attentional resources can be allocated to other processes to raise the quality of comprehension. From a capacity-limited cognitive process introduced above, it follows that it is indispensable to develop fluent or automatic word recognition and syntactic processing skills for the accomplishment of fluent reading, which is the key indicator of a highly skilled reader.

### Previous Research on L1 Reading Fluency

Research on L1 reading fluency can be divided into two major categories: group and individual comparison studies and training studies. The comparison studies involve fluency research on word recognition skills, oral text reading, enjoyment of reading, and reading rates. Many L1 comparison studies have documented the impact of word recognition speed and accuracy on reading comprehension. On the other hand, the great majority of training studies on L1 reading fluency involve various types of repeated reading practices, such as "(a) repeated-reading practice of a text with a partner; (b) radio reading; (c) 'say it like a character;' (d) mumble reading; (e) cooperative reading; and (f) reading-center practice with an audiotape (Rasinski, 2003). Surprisingly, there are relatively few studies considering the positive impact of extensive reading or reading rate in the L1 reading research literature. The present study examines the previous studies in terms of the impact of word recognition and extensive reading on reading fluency. Syntactic processing fluency is not examined here, because children reading in L1 have already acquired grammatical knowledge implicitly when they enter elementary schools.

*Word-recognition Fluency*

Many studies comparing individual and group differences among elementary school students have demonstrated that word recognition skills are strongly associated with reading comprehension. However, "beyond third-grade level, word recognition skills do not usually predict reading comprehension for better readers" (Grabe, 2009, p. 301). Most L1 evidence indicates a sharp drop in the relationship between word recognition and comprehension from third-grade level (Perfetty, 2007). As opposed to good readers, weaker readers generally demonstrate a strong relationship between word recognition skills and reading through adulthood (Stanovich, 2000). One important recent study has demonstrated the power of word recognition fluency, even for fifth-grade students. Klauda and Guthrie (2008) assessed 278 fifth-graders on word recognition fluency, syntactic fluency and passage reading fluency in relation to reading comprehension. Word recognition fluency, on its own, accounted for 43% of the shared variance with reading.

With respect to word recognition training or practice, a few studies have shown that the practice not only leads to faster word recognition but also improves reading comprehension when words trained on appear in a reading text. However, most word recognition practice studies do not show a significant positive impact directly on reading comprehension (Grabe, 2010, p. 75).

*Effects of Extensive Reading on Reading Fluency*

Research on passage reading practice studies has had a significant impact on reading comprehension understanding and training. In particular, training that involves students engaging in repeated reading of texts almost uniformly led to improved comprehension outcomes (National Reading Panel, 2002). Kuhn and Stahl (2003) argued that the positive results associated with repeated reading practices may be associated more with elementary-level students and with reading disability students. Concerning extensive reading, there are few L1 studies that specifically address the impact of extensive reading on reading comprehension improvement. In one such study, Homan, Klesius, and Hite (1993) demonstrated that a 7-week extensive reading practice as well as repeated reading among 6th-grade L1 students led to significant gains as seen in a pre-test and post-test comparison. Overall, training studies to demonstrate strong linkages between comprehension abilities and extensive reading appear extremely limited in L1 contexts.

## Previous Research on L2 Reading Fluency

ESL or EFL teachers, particularly in Japan, have seen accuracy and fluency as distinct and potentially contrasting skills in language development. Fluency has generally been ignored, as in intensive reading we usually read at a slower speed and are required to gain a higher degree of accurate understanding. However, a strong psychological perspective by Schmidt (1992) and Segalowitz and Segalowitz (1993) has explained that fluency and accuracy are not competing factors in language performance, that fluency promotes accuracy, and that accuracy is an indication of increasing fluency in language performance. More and more pioneer EFL teachers have recognized the importance of fluency development for language learning, especially for reading, and have incorporated extensive reading in their curricula.

*Word-recognition Fluency*

A few L2 comparison studies have examined the impact of word recognition skills on reading comprehension. Verhoeven (2000) has demonstrated a significant causal relationship between word recognition skills and reading comprehension. Turning to L2 word-recognition fluency, Akamatsu (2008), Segalowitz and Segalowitz (1993), and Fukkink, Hulstijn and Simis (2005) have all investigated whether direct word-recognition training will develop word recognition abilities. Their results showed that word-recognition training improved students' word recognition performance in both speed and accuracy. In the case of Fukkink, Hulstijn and Simis, a significant improvement in reading comprehension was not found after 2 days of word recognition training. Overall, fluent word recognition skills play an important role in reading comprehension, but it may safely be said that word-recognition training does not seem to improve reading comprehension.

*Syntactic Processing Fluency*

Previous studies showed that grammatical knowledge is closely associated with reading comprehension. Shiotsu and Weir (2007, p. 128) showed that the correlation between reading and syntactic knowledge was $r$=.85. Judging from the fact that they didn't measure response time in the syntax test, this syntactic knowledge does not demonstrate syntactic processing skills. On the other hand, Nassaji (1998) figured out efficiency scores and showed that the correlation between syntactic processing skills and reading comprehension in terms of efficiency scores was $r$= - .51. (Efficiency scores in his study were computed based on the following formula: efficiency scores = ($Z$-scores of errors＋$Z$-scores of speed ) / 2.

*Effects of Extensive Reading on Reading Fluency*

A few passage rereading fluency training studies have been conducted in the past 15 years, and have provided indications as to improvements in reading fluency, and in some cases, in reading comprehension. However, since rereading tasks are very monotonous for normal students, extensive reading is preferable to be incorporated into curricula. As for extensive reading and its impact on reading comprehension and reading fluency, several studies are reviewed in Grabe (2009), Iwahori (2008), Fujita and Noro (2009) and others.

*Research Questions*

Overall, the L2 reading fluency research, while limited in the number of studies, generally supports the importance of word reading fluency, passage reading fluency, extensive reading and reading rate training on reading comprehension improvement. However, to the best of our knowledge, there has been hardly any research done as to whether extensive reading will improve reading comprehension ability, reading fluency, word recognition fluency or syntactic processing fluency. Based on the previous studies, the following research questions (RQs) were addressed in this study.

RQ1: Can extensive reading improve the speed of as well as the accuracy on reading comprehension tests and cloze tests?
RQ2: Can extensive reading improve the speed of as well as the accuracy of word recognition?
RQ3: Can extensive reading improve the speed of as well as the accuracy of syntactic processing?
RQ4: What contribution do word recognition and syntactic processing make to the results of reading comprehension tests and cloze tests?

## Method

### Participants

120 first-year students in the Department of Global English were involved in this program. However, those who missed one of the four types of tests (See 'Instruments' in the following section.) conducted in April as a pretest and in December as a post-test were excluded from this study. As a result, the data were collected from 105 students. Extensive Reading is a required course for the students in the Department of Global English. Students were expected to borrow graded readers from the library, to read as many books as possible and to submit a report on the book or books read during the previous week out of class. In class, students were engaged in (1) *sustained silent reading* while the instructors were commenting on a book report individually, 2) *reading-while-listening practice* so as to be able to read without regression to the previous part of a sentence or text, and (3) *story retelling* pair work to improve speaking skills through the introduction of the stories they read.

## Instruments

The following tests were conducted as pre-tests at the beginning and as post-tests at the end of about 7 months' extensive reading practice: placement/progress tests (cloze tests), extensive reading tests, computer-based English lexical processing tests and syntactic processing tests (grammaticality judgment tests).

### *Placement/progress tests (PPTs)*

E.P.E.R. (Edinburgh Project on Extensive Reading) Placement Test A was conducted. PPTs are cloze tests, each of which is composed of more than 10 graded passages of increasing difficulty. There are three parallel versions available (Version A, Version B and Version E, though Version A seems to be easiest), of which Version A was used in the present study. There are 141 blanks in Version A. Each correct answer is scored 1. PPTs are not strictly fixed ratio cloze tests (namely, every *n*th word is not deleted) but semi-fixed ratio tests. PPTs are scored according to the Answer Key provided by EPER, in which only one answer is given to many blanks but some blanks can accommodate alternative lexical or grammatical variations. Words which are incorrectly spelt are accepted if they are semantically acceptable. It can be said that PPTs are basically scored by an exact-word method, but try to combine an acceptable-word method within this limited range. (EPER, 1992)

The same test (Version A) was administered as a pre-test and as a post-test in this study. Unlike reading comprehension tests, cloze tests seem to have fewer test-retest effects, because the tests consist of short passages, and lay more emphasis on lower-level linguistic competence, so that the students seemed not to remember the content of the test after a 7-month treatment.

### *Extensive Reading Tests (ERTs)*

ERTs are original or unique in that they aim to assess *extensive reading ability*. In many cases with extensive reading research, to assess the development of reading ability after an extensive reading program, TOEFL, TOEIC or STEP style reading tests are constructed mostly based on short expository texts. In the light of the definition of extensive reading, these reading tests do not reflect reading quantity, reading speed, or reading materials that arouse the interest of readers and match the student's proficiency. Since ERTs aim at assessing extensive reading ability or general reading ability, the texts for ERTs are taken from English graded readers and are relatively long (Level E Version 1 - approximately 1,260 words with 18 questions; Level F Version 2 - approximately 950 words with 22 questions). Most of the questions are global comprehension questions. There are eight levels in ERTs (X, A, B, C, D, E, F, G, from the advanced to the elementary). Various types of comprehension questions are included: yes/no questions, multiple-choice questions, questions demanding short answers in English, gap-fill or gap-summary questions.

### *Computer-Based English Lexical Processing (CELP) tests*

Schmitt (2010, p. 242) referred to "speed of processing" as an important part of vocabulary knowledge because it needs to be used effectively in real-life L2 communication. Moreover, Daller et al. (2007, p. 8) explained vocabulary competence utilizing three aspects, adding a processing component named "fluency" to "breadth" and "depth". To be sure, it is significant to add lexical access speed as a new component, but the reliability has not been investigated. In order to validate lexical access speed as one component of vocabulary ability, Kadota et al. (2015) developed the CELP test, which aims to measure both the accuracy rate (how many words can be answered correctly) and reaction time (lexical access speed) of the given list of words. The vocabulary list used by the CELP test consists of 100 items including synonym and non-synonym pairs, some of which are shown in Table 1. The CELP test uses a synonym judgment test based on the cognitive theory of a priming effect to directly measure lexical access speed on the top of accuracy. The reliability of the test has been investigated: $\alpha$ =.89 (Kadota, 2010a, p. 82), and correlation has been found with other established standardized vocabulary tests, such as the Vocabulary Size Test (Nation and Beglar, 2007): $r = .64$ (Kadota, 2010b).

One hundred items (50 synonym pairs; 50 non-synonym pairs) were presented to have test-takers judge whether each pair of target word and prime word fell into a synonym pair or not, as shown in Figure 1. The scores of the CELP test were shown as the number of correct responses (CELP-A: accuracy rate) and the reaction time of correct responses (CELP-R: lexical access speed).

Table 1

*Some Examples of Synonym and Non-synonym Pairs Used in CELP-test (A Version)*

| Synonym pairs | | Non-synonym pairs | |
|---|---|---|---|
| **Prime words** | **Target words** | **Prime words** | **Target words** |
| get | acquire | knock | adjust |
| enough | sufficient | seat | anger |
| find | discover | strange | pleasant |
| chance | opportunity | design | struggle |
| fast | quick | action | wave |
| show | display | create | release |
| choice | selection | dirty | precise |
| particular | specific | energy | wave |
| problem | trouble | happy | dry |
| idea | concept | poem | fortune |

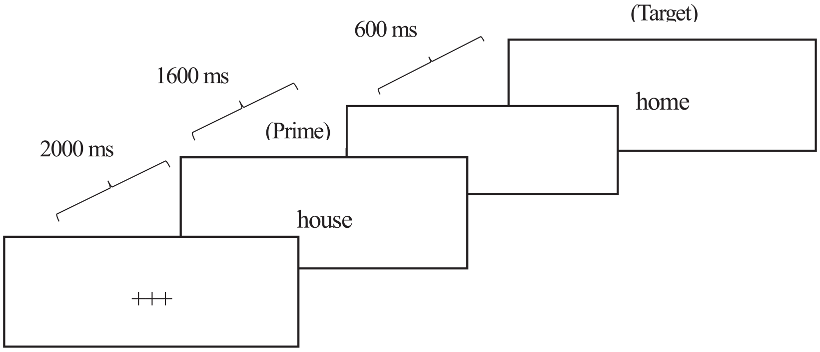


*Figure 1.* Procedure of the CELP test used in the pre- and post-tests. Computer screens are represented by white boxes. +++ depict gazing points. After +++ are gazed at for 2000 milliseconds, a prime word (e.g. house) is presented and disappears in 1600 milliseconds. 600 milliseconds after the disappearance, a target word (e.g. home) pops out on the display. After the target is presented, the participants are required to press "B" for YES (synonym) or "N" for NO (non-synonym).

*Syntactic Processing Tests (Grammaticality Judgment Tests: GJTs)*

In addition to word recognition, another lower-level process is syntactic processing. A fluent reader is able to take in and store words together so that basic grammatical information can be extracted to support clause-level meaning. To investigate how to process a text syntactically in reading comprehension, the present writers conducted a timed grammaticality judgment test (GJT) in this study. A GJT can be defined as a task in which participants are involved in "deciding whether a sentence is well-formed or deviant" (Ellis, 1991, p. 162). The question has arisen as to whether GJTs provide a direct measure of linguistic competence or a performance measure of L2 learners' linguistic abilities. "The important issue is to determine the nature of linguistic knowledge, whether implicit, explicit or some combination of both, that participants draw on in their performance of GJTs" (Loewen, 2009, p. 95).　If learners are given limited time to respond, this may encourage them to rely on implicit knowledge; in contrast, unlimited time may allow them to access explicit knowledge (Bialystok, 1979). The reason why timed GJTs were employed in this study is that it is likely that grammatical processing skills are learned implicitly, as mentioned in the literature review, and that a rapid response time (or latency) may indicate a high degree of availability and automaticity of the required grammatical knowledge (Alanen, cited in Loewen, 2009).

Table 2

*Some Examples of Grammatical Sentences and Ungrammatical Sentences Used in Grammaticality Judgment Tests*

| Grammatical Sentences ( YES responses ) | No. of words | Ungrammatical Sentence ( NO responses ) | No. of words | Target grammar | Domain of grammar | Points to judge |
|---|---|---|---|---|---|---|
| She bought her son a skateboard as a birthday present. | 10 | He bought a handbag his girl friend at the store. | 10 | indirect object | word order | 5 |
| Bill wanted to know where I had been. | 8 | Betty wanted to know why had he studied Chinese. | 9 | indirect questions | word order | 6 |
| Taro works very hard but earns very little. | 8 | Hanako lives in Kyoto but work in Osaka. | 8 | third person -s | morphology | 6 |
| Hiroshi found some keys on the ground. | 7 | Hanako bought two present for her children. | 7 | plural -s | morphology | 4 |
| He is nicer and more intelligent than the other students. | 10 | My car is more faster and more powerful than yours. | 10 | comparison | sentence & phrase | 5 |
| The car that Bill has rented is a Toyota. | 9 | The boat that Tom bought it has been stolen. | 9 | relative clauses | sentence & phrase | 6 |
| She arranged the flowers she had picked in the garden. | 10 | I showed him the picture I took the day before. | 10 | perfect form | aspect | 7 |

With reference to the GJT items used in Loewen's study (2009, pp. 356-358), the GJT was made for this study. (Some examples of grammatical and ungrammatical items are listed in Table 2.) The test items were controlled in terms of sentence length (7 to 10 words) and points in judging whether grammatical or ungrammatical sentences ($4^{th}$ to $7^{th}$ word). The test consists of 80 sentences, covering 20 grammatical structures. For each of the 20 structures, two grammatical and two ungrammatical sentences were included in the GJT, resulting in 80 items. Six structures are concerned with morphological knowledge, such as past tense form, plural –s, third person –s, declension of pronouns, articles, and auxiliary verbs. Fourteen structures are concerned with syntactical knowledge, such as word order (five basic sentence structures, indirect questions, participles, adverb placement, and prepositional clauses), phrases and sentences (infinitives, gerunds, participles, comparisons, tag-questions and relative clauses), and tense and aspect (perfect tense, progressive forms, and passive voice). The timed GJT, which was written in a programming language called HSP (Hot Soup Processor), is also a computer-based test. The GJT was delivered to the students' computers to have them run the test individually. When the target is presented, the participants are required to press "B" for YES (grammatical) or "N" for NO (ungrammatical) as quickly as possible.

**Research Design**

A pre/post-test design was used for the experimental group in this study. It is ideal to have a few control groups in an experimental study (Nakanishi, 2015), to be sure, but it is impossible to set a control group, because all the participants in the Department of Global English are involved in the Extensive Reading course as a required subject. Thus, it is very difficult to create true experimental conditions in real educational settings (Grabe, 2004).

A counterbalance design was adopted for ERTs. ERT Level E was used for A and B classes, ERT Level F for C and D classes. Each level has two versions: Version 1 and Version 2. It is inappropriate to use the same reading comprehension test both as a pre-test and as a post-test, because the content of a long narrative text can be remembered for some time. As Table 2 shows, Version 1 and Version 2 were counterbalanced: half of A & B class participants took Version 1 for the pre-test, and the other half had Version 2, and vice versa for the post-test.

(1) Pre- and post-tests were conducted at the beginning and at the end of the extensive reading program.
(2) The extensive reading program continued for about 7 months. (There was a summer vacation of approximately 45 days between the two terms).
(3) The participants were divided into A and B groups and C and D groups, and given different levels of ERTs.
(4) A control group could not be set, because all the participants belonged to the same department, and extensive reading was one of their required subjects.

Table 2
*Counterbalanced Measures Design for Version 1 and Version 2 of EPER Extensive Reading Tests*

| | Pretest | Posttest |
|---|---|---|
| Group 1 of A & B classes | E-level Version 1 | E-level Version 2 |
| Group 2 of A & B classes | E-level Version 2 | E-level Version 1 |
| Group 1 of C & D classes | F-level Version 1 | F-level Version 2 |
| Group 2 of C & D classes | F-level Version 2 | F-level Version 2 |

**Procedure**

The participants had four types of tests (PPTs, ERTs, CELP tests, & GJT tests) as pre-tests in April and as post-tests in December. PPTs (Placement/Progressive Tests) and ERTs (Extensive Reading Tests) were conducted in the classrooms. The CELP tests and GJT tests were administered in a computer room. The application of CELP test and GJT was distributed to the students' computers to have them run the test individually. All the procedures of those tests were explained on the computer screens they were looking at and the participants carried out the tests on their own.

Performance on the experimental tasks was measured in terms of scores for both accuracy and speed (reaction time or latency). Time was measured because simply knowing how to perform a task may not ensure that a reader is able to process the task components as rapidly and accurately as required during fluent reading. An index based on both speed and accuracy can provide a more accurate measurement of individual differences in reading than accuracy alone. Time for ERTs was measured in seconds by the clock in the classroom. However, as to PPTs, all the participants were given fifteen minutes during which time they were required to fill in as many blanks as possible. CELP tests and GJTs were programmed so as to measure response time in milliseconds. The two indices of speed and accuracy were combined to yield indices of the reader's efficiency in processing tasks. Efficiency scores were computed based on the formula: accuracy scores divided by response time.

## Results and Discussion

Table 4 shows the accuracy scores and standard deviation of the pre- and post-tests on reading progress tests, cloze tests, word recognition tests and syntactic processing tests. The score differences between pre-tests and post-tests indicate the steady development of reading comprehension in A and B classes, and cloze skills. On the contrary, almost no improvement can be found in reading comprehension C and D group, word recognition, and syntactic processing. It is very interesting to note the different results between the YES response items and NO response items, and between grammatical items and ungrammatical items. Concerning the Word Recognition Test, the participants can judge more correctly NO response items as non-synonym pairs than the YES response items as synonym pairs. Similarly, in the Syntactic Processing Test, the subjects seem to find it more difficult to judge Grammatical Items as correct sentences than Ungrammatical Items as incorrect sentences.

Table 4

*Mean **Accuracy Scores** and Standard Deviations of the Pre- and Post-Tests*

| Variables | | | Pretest | | Posttest | |
|---|---|---|---|---|---|---|
| | MPS | *n* | *M* | *SD* | *M* | *SD* |
| Reading Progress Test | | | | | | |
| E-level Test (A & B Classes) | 28 | 55 | 17.76 | 4.910 | 20.98 | 3.546 |
| F-level Test (C & D Classes) | 28 | 50 | 16.24 | 5.855 | 17.44 | 5.729 |
| Cloze Tests | | 105 | 29.31 | 11.356 | 35.33 | 12.175 |
| Word Recognition Test | 100 | 105 | 79.99 | 8.232 | 80.63 | 7.823 |
| YES Response Items | 50 | 105 | 36.07 | 5.460 | 35.73 | 5.723 |
| NO Response Items | 50 | 105 | 43.86 | 4.735 | 44.90 | 4.081 |
| Syntactic Processing Test | 80 | 105 | 48.76 | 7.608 | 49.48 | 6.905 |
| Grammatical Items | 40 | 105 | 29.60 | 4.922 | 29.47 | 5.071 |
| Ungrammatical Items | 40 | 105 | 19.16 | 5.500 | 20.01 | 4.817 |

*Note*. MPS= Maximum Possible Scores

Table 5 shows the response times and standard deviation of the pre- and post-tests on reading progress tests, word recognition tests and syntactic processing tests. From the fact that the response times of the three tests in the post-test became shorter than those in the pre-test, it is obvious that processing skills turned out to be faster. It is important to note that students find it more difficult to recognize nonsynonym pairs correctly than synonym pairs. Contrary to this tendency, the results of the Syntactic Processing Tests indicate that it is more difficult to judge grammatical sentences as grammatically correct sentences than ungrammatical sentences. It is assumed that in processing ungrammatical items and finding ungrammatical points, they respond immediately.　In the case of grammatical items, however, it can be inferred that they process the sentences again and again trying to find grammatically incorrect parts until they are confident of no mistake in a target sentence.

Table 5

*Mean* ***Response Time****s and Standard Deviations of the Pre- and Post- Tests*

| Variables | | | Pretest | | Posttest | |
|---|---|---|---|---|---|---|
| | MPS | *n* | *M* | *SD* | *M* | *SD* |
| Reading Progress Test (min.) | | | | | | |
| E-level Test (A & B Classes) | 28 | 55 | 25.726 | 3.866 | 22.539 | 5.462 |
| F-level Test (C & D Classes) | 28 | 50 | 26.94 | 3.661 | 23.04 | 4.567 |
| Word Recognition Test (sec.) | 100 | 105 | 107.850 | 33.897 | 92.846 | 28.458 |
| YES Response Items | 50 | 105 | 46.257 | 15.297 | 39.452 | 14.644 |
| NO Response Items | 50 | 105 | 61.503 | 20.456 | 53.459 | 16.216 |
| Syntactic Processing Test (sec.) | 80 | 105 | 347.483 | 125.926 | 320.042 | 117.393 |
| Grammatical Items | 40 | 105 | 205.699 | 71.616 | 186.396 | 66.329 |
| Ungrammatical Items | 40 | 105 | 141.783 | 64.274 | 133.645 | 58.468 |

*Note*.　MPS= Maximum Possible Scores; min. = minutes; sec. =seconds

Table 6 indicates the mean efficiency scores and standard deviations of the pre- and post-tests on four types of tests. The efficiency scores were computed based on the formula: accuracy scores divided by response times (minutes). In the four types of tests, average efficiency scores increased. Namely, the scores, on the whole, prove that they came to process words, sentences, and grammar faster and more accurately after 7 months' extensive reading practice.

Table 6

*Mean* ***Efficiency Scores*** *and Standard Deviations of the Pre- and Post- Tests*

| Variables | | Pretest | | Posttest | |
|---|---|---|---|---|---|
| | *n* | *M* | *SD* | *M* | *SD* |
| Reading Progress Test (accuracy scores / min.) | | | | | |
| E-level Test (A & B Classes) | 55 | .737 | .291 | 1.001 | .379 |
| F-level Test (C & D Classes) | 50 | .625 | .268 | .784 | .286 |
| Cloze Tests (accuracy scores / 15 min.) | 105 | 1.954 | .757 | 2.355 | .811 |
| Word Recognition Test (accuracy scores / min.) | 105 | 27.792 | 31.197 | 31.267 | 34.410 |
| YES Response Items | 105 | 13.021 | 14.282 | 14.576 | 15.400 |
| NO Response Items | 105 | 15.440 | 17.235 | 17.662 | 19.251 |
| Syntactic Processing Test (accuracy scores /min.) | 105 | .156 | .053 | .180 | 093 |
| Grammatical Items | 105 | .158 | .050 | .182 | .091 |
| Ungrammatical Items | 105 | .154 | .070 | .180 | .110 |

Table 7 shows *t*-test results between average scores of pre- and post-tests on reading comprehension tests, cloze tests, word recognition tests, and syntactic processing tests. The results display differences in the averages of most tests, except the accuracy scores of the Reading Progress Test (C and D group), the Word Recognition Test, and the Syntactic Processing Test. In the three tests, the mean scores of the post-tests were a little better than those of the pre-tests, but they didn't turn out to be statistically significant. As for the Reading Progress Test (C and D group), the participants' attitude to reading accurately or fast might influence reading comprehension. Those who want to read the text correctly take more time than is necessary, without minding reading time, and vice versa about fast readers. Concerning the Word Recognition Test and the Syntactic Processing Test, the learners might use their attention or resources of working memory to grasping the meaning of the stories, so that their resources might not be left for learning new vocabulary and grammatical features incidentally. Besides, it is important to point out that the graded readers they read are not supposed to contain many unknown words and new grammatical features. Cohen's *d* indicates effect size (the magnitude of the difference between two groups). Most indices of effect size correspond to those of the *t*-tests. However, the Cohen's *d* index of word recognition (efficiency scores) is quite small (-.10), even though the *t*-test displays statistical difference (-3.092, $p < .01$). The results of accuracy scores and latency of the CELP test might not have been stable. Cohen's *d* index (-.24) of reading comprehension in C and D group is not as small as the *t*-vale indicates ($t$= -1.235). This difference is partly caused by the differences of *SD* scores between C and D group and A and B group (See Table 4).

Table 7

*T-tests Between Pre- and Post-Tests on Reading Comprehension, Cloze, Word Recognition, and Syntactic Processing*

| Variables | *n* | *M* | *SD* | *SE* | *t*-value | Cohen's *d* |
|---|---|---|---|---|---|---|
| Reading Progress Pre- and Post-Tests | | | | | | |
| E-level Test (Accuracy) | 55 | -3.218 | 3.914 | .528 | -6.097*** | -.75 |
| E-level Test (Response Time) | 55 | 2.736 | 3.661 | .493 | 5.542*** | .57 |
| E-level Test (Efficiency) | 55 | -.273 | .244 | .032 | -8.315*** | -.80 |
| F-level Test (Accuracy) | 50 | -1.200 | 6.872 | .972 | -1.235 | -.24 |
| F-level Test (Response Time) | 50 | 3.900 | 4.837 | .684 | 5.701*** | .94 |
| F-level Test (Efficiency) | 50 | -.158 | .312 | .044 | -3.582** | -.57 |
| Cloze Pre- and Post-Tests | | | | | | |
| Accuracy Scores | 105 | -6.016 | 8.74 | .854 | -7.05*** | -.51 |
| Efficiency Scores | 105 | -.401 | .583 | .056 | -7.05*** | -.51 |
| Word Recognition Pre- and Post-Tests | | | | | | |
| Accuracy Scores | 105 | -.638 | 4.42 | .432 | -1.476 | -.07 |
| Response Times | 105 | 15.00 | 28.93 | 2.82 | 5.313*** | .62 |
| Efficiency Scores | 105 | -3.475 | 11.518 | 1.124 | -3.092** | -.10 |
| Syntactic Processing Pre- and Post-Tests | | | | | | |
| Accuracy Scores | 105 | -.714 | 5.566 | .543 | -1.315 | -.05 |
| Response Times | 105 | 27.441 | 113.473 | 11.073 | 2.478* | .57 |
| Efficiency Scores | 105 | -.0248 | .0864 | .0084 | -2.939** | -.72 |

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

In order to determine the relative contribution of the predictor variables to the criterion variables, a series of forced-entry multiple regression analyses were performed based on the results of the four types of post-tests, which are the outcomes of seven months' extensive reading practice. After the forced-entry regression analyses, one of the two predictor variables was excluded. However, the coefficient of determination ($R^2$) of the two predictor variables was a little larger than that of one significant predictor variable, so stepwise regression analyses were performed to clarify $R^2$ change in the second model. Table 8, 9, 10, 11, 12, and 13 display the results of six hierarchical multiple regression analyses where the predictor variables are the accuracy and the efficiency scores on the component measures (word recognition and syntactic processing) and the criterion variables are the accuracy and the efficiency scores on cloze and reading comprehension. The column labeled $R^2$ change in each table gives the magnitude of the contribution of each variable at the point where the variable was entered into the analysis. As Table 8 shows, accuracy on the syntactic processing measure accounted for 19.2 % ($p < .01$) of the variance in cloze measure. When word recognition measure was entered into model 2, it added 2.2 % of the variance in cloze, but it did not reach significance. Therefore, it may safely be said that syntactic processing significantly explained 19.2 % of the variance in cloze measure. On the other hand, as Table 9 shows, efficiency scores on the word recognition measure accounted for 13% ($p < .001$) of the variance. When word recognition was entered into the regression model, the second model accounted for an additional 2.2 % of the variance, but word recognition measure did not contribute to cloze performance. It can be concluded that regression analyses using efficiency scores showed that only syntactic processing measure was a significant predictor in cloze measure.

Table 8

*Hierarchical Multiple Regression Using Accuracy Scores on Component Measures (Word Recognition and Syntactic Processing) as Predictor Variables and Cloze as the Criterion Variable* ($N$=105)

| Predictor | *B* | *SE B* | $\beta$ | *p* | $R^2$ (Adjusted $R^2$) | $R^2$ change |
|---|---|---|---|---|---|---|
| 1 | | | | | .192 (.117) | .192 |
| Syntactic Processing | .773 | .156 | .438 | .000*** | | |
| 2 | | | | | .214 (.193) | .022 |
| Syntactic Processing, | .616 | .182 | .349 | .001** | | |
| Word Recognition | .267 | .160 | .172 | .099 | | |

*Note*. ** $p < .01$, *** $p < .001$

Table 9

*Hierarchical Multiple Regression Using Efficiency Scores on Component Measures (Word Recognition and Syntactic Processing) as Predictor Variables and Cloze as the Criterion Variable* ($N$= 105)

| Predictor | *B* | *SE B* | $\beta$ | p | $R^2$ (Adjusted $R^2$) | $R^2$ change |
|---|---|---|---|---|---|---|
| 1 | | | | | .130 (.122) | .130 |
| Word Recognition | -.009 | .002 | -.361 | .000*** | | |
| 2 | | | | | .137 (.120) | .007 |
| Word Recognition | -.009 | .002 | -.380 | .000*** | | |
| Syntactic Processing | -.734 | .815 | .085 | .370 | | |

*Note*. *** $p < .001$

Table 10 and Table 11 show the results of two hierarchical multiple regression analyses using accuracy (Table 10) and efficiency scores (Table 11) on component measures of E-level PPT users in A and B classes. The criterion variable is reading comprehension. As Table 10 shows, accuracy on syntactic processing measure accounted for 17 % ($p < .01$) of the variance in reading comprehension measure; however, word recognition measure, when syntactic processing measure was entered into the model 2, accounted for a very small significant additional variance (2.5 % of the variance) in reading comprehension, but syntactic processing measure did not reach significance. On the other hand, as Table 11 shows, efficiency on word recognition measure accounted for 10.8 % of the variance in reading comprehension, but it did not reach statistical significance, when syntactic processing measure was added into the model 2. However, efficiency on syntactic processing measure added a significant variance (8.5 % of the variance) to reading performance,

Table 10
*Hierarchical Multiple Regression Using Accuracy Scores on Component Measures (Word Recognition and Syntactic Processing) as Predictor Variables and Reading Comprehension as the Criterion Variable* (A & B classes; *N*=55)

| Predictor | *B* | *SE B* | $\beta$ | *p* | $R^2$ (Adjusted $R^2$) | $R^2$ change |
|---|---|---|---|---|---|---|
| 1 | | | | | .170 (.154) | .170 |
| Word Recognition | .334 | .102 | .412 | .002** | | |
| 2 | | | | | .195 (.164) | .025 |
| Word Recognition | .318 | .102 | .392 | .003** | | |
| Syntactic Processing | .095 | .075 | .158 | .212 | | |

*Note*. ** $p < .01$

Table11
*Hierarchical Multiple Regression Using Efficiency Scores on Component Measures (Word Recognition and Syntactic Processing)as Predictor Variables and Reading Comprehension as the Criterion Variable* (A & B classes; *N*=55)

| Predictor | *B* | *SE B* | $\beta$ | *p* | $R^2$ (Adjusted $R^2$) | $R^2$ change |
|---|---|---|---|---|---|---|
| 1 | | | | | .108 (.091) | .108 |
| Word Recognition | .478 | .188 | .329 | .014* | | |
| 2 | | | | | .193 (.162) | .085 |
| Word Recognition | .325 | .192 | .224 | .097 | | |
| Syntactic Processing | 2.390 | 1.020 | .310 | .023* | | |

*Note*. * $p < .05$

Table 12 and Table 13 show the results of two hierarchical multiple regression analyses using accuracy (Table 12) and efficiency scores (Table 13) on component measures of F-level PPT users in C and D classes. The criterion variable is reading comprehension. As Table 12 shows, accuracy on word recognition measure accounted for 1.3 % of the variance and accuracy on syntactic processing measure explained 4.7% of variance in reading comprehension measure, but neither were statistically significant. In the same manner, as Table 13 shows, efficiency on word recognition measure accounted for 1.8 %, but it did not reach significance. Efficiency on syntactic processing explained 6.7 %, but it did not show any statistical significance. From the results of regression analyses, it follows that accuracy and efficiency scores on word recognition and syntactic processing measures in the C and D groups did not contribute to explaining the statistically significant variance of reading comprehension skills.

Table 12

*Hierarchical Multiple Regression Using Accuracy Scores on Component Measures (Word Recognition and Syntactic Processing) as Predictor Variables and Reading Comprehension as the Criterion Variable* (C &D classes: *N*=50)

| Predictor | *B* | *SE B* | *β* | *p* | $R^2$ (Adjusted $R^2$) | $R^2$ change |
|---|---|---|---|---|---|---|
| 1 | | | | | .169 (.152) | .169 |
| Word Recognition | .357 | .114 | .412 | .003** | | |
| 2 | | | | | .182 (.147) | .012 |
| Word Recognition | .326 | .120 | .377 | .009** | | |
| Syntactic Processing | .119 | .142 | .116 | .407 | | |

*Note*. ** $p < .01$

Table 13

*Hierarchical Multiple Regression Using Efficiency Scores on Component Measures (Word Recognition and Syntactic Processing) as Predictor Variables and Reading Comprehension as the Criterion Variable* (C&D classes: *N*=50)

| Predictor | *B* | *SE B* | *β* | *p* | $R^2$ (Adjusted $R^2$) | $R^2$ change |
|---|---|---|---|---|---|---|
| 1 | | | | | .064 (-.044) | .064 |
| Syntactic Processing | -.588 | .325 | -.253 | .077 | | |
| 2 | | | | | .085 (.046) | .067 |
| Word Recognition | .002 | .002 | .146 | .303 | | |
| Syntactic Processing | -.603 | .325 | -.259 | .070 | | |

Over all, in step 1 of the analyses except Table 13, accuracy or efficiency scores on a predictor variable accounted for statistically significant variance showing the contribution of word recognition or syntactic processing measures to cloze or reading comprehension measures. Judging from the fact that $R^2$ (coefficient of determination) is figured out by squaring $r$ (correlation coefficient), there should be a medium or high correlation coefficient between the variables. Actually, however, it is assumed that correlations between predictor variables (such as word recognition and syntactic processing) and criterion variables (such as cloze or reading comprehension measures) were inconsistent or unstable. Accuracy scores on cloze and reading comprehension measures developed, but accuracy scores on word recognition and syntactic processing remained unchanged.

## Conclusion

This study attempted to examine the development of cloze skills, reading comprehension ability, word recognition and syntactic processing skills through extensive reading practice, and the contribution of word recognition and syntactic processing to cloze and reading comprehension. As for RQ1, as *t*-tests in Table 7 show, almost all the accuracy scores on cloze measure and reading comprehension measure developed significantly, with the exception of the accuracy scores of reading comprehension in C and D class. Response times and efficiency scores of cloze skills and reading comprehension improved significantly. As to RQ2 and RQ3, as *t*-tests in Table 7 indicate, the accuracy scores of word recognition and syntactic processing did not develop, but the processing speed of the component skills improved so that the efficiency scores on word recognition and syntactic processing measures also developed significantly. One of the reasons why the accuracy scores did not develop is that the participants who are freshmen without any experience in extensive reading have used up their cognitive resources in working memory to grasp the meaning of the texts, so that their resources might not be sufficient for learning new linguistic items. Another reason is that most graded readers they read are written in easy English - comprehensible words and grammatical structures. Therefore, they do not come across new words and structures so frequently. The other reason is that grammatical judgment tests were used for measuring syntactic processing skills in this study, but grammatical knowledge might be essentially different from syntactic parsing skills for word ordering, subordinate clauses and pronominal forms, and for isolating the main verb, subject position, object position and any adjunct noun-phrase positions. It seems to be indispensable that the grammatical judgment test must be revised so as to be able to measure syntactic processing skills effectively.

The response time was shortened between pre-tests and post-tests in the four different types of tests, which means that the processing skills in all the tests have become efficient. At the beginning the learners read word by word translating English into Japanese slowly. However, through the exposure to a lot of easy English words and phrases they have come to identify and recognize them faster, and to read more rapidly little by little without translating into Japanese, and without regressions to previous parts of the sentence or text. From such processing improvements, it can be safely be said that readers develop processing skills first, and then develop reading skills and linguistic knowledge. It is often pointed out by many reading researchers that we are required to read 1,000,000 words to become fluent readers. It follows that our students need one or two more years to complete the reading target of 1,000,000 words before becoming fluent readers.

Concerning RQ4, the multiple regression analyses demonstrated the contribution of component skills to cloze and reading comprehension to some extent. The analyses displayed that the efficiency scores on word recognition accounted for about 14% in cloze measures, while the accuracy scores on syntactic processing explained 19% of variance in cloze measure. The accuracy scores on word recognition measure accounted 13% of variance in reading, while the results of the C & D group failed to show the contribution of component skills to reading measure. The tests for measuring response time in milliseconds can easily be affected by various factors such as participants' motivation, proficiency, and attitude to the tests. When the testees in a lower–level class asked for the post-tests if the results of these tests would be included in the evaluation, the teacher said no, and their eagerness or concentration seemed to lapse. It is necessary to explore the contribution of word recognition and syntactic processing to reading comprehension again, taking such participants' mental state into consideration.

## References


Akamatsu, N. (2008). The effects of training on automatization of word recognition in English as a foreign language. *Applied Psycholinguistics, 29*, 175-193.

Alanen, R. (1995). Input enhancement and rule presentation in second language acquisition. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 259-302). Honolulu, HI: University of Hawai'i.

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

Bialystok, E. (1979). Explicit and implicit judgments of L2 grammaticality. *Language Learning, 29*, 81-103.

Breznitz, A. (2006). *Fluency in reading: Synchronization of processes*. Mahwah, NJ: Laurence Erlbaum.

Car, T., & Levy, B. (Eds.) (1990). *Reading and its development: Component skills approaches*. San Diego: Academic Press.

Carver, R. (1997). Reading for one second, one minute, or one year from the perspective of reading theory. *Scientific Studies of Reading*, *1*, 3-43.

Daller, H., Milton, J., & Treffers-Daller, J. (2007). *Modelling and assessing vocabulary knowledge*. New York:

Cambridge University Press.

EPER (Edinburgh Project on Extensive Reading). (1992). *The EPER guide to organizing programmes of extensive reading*. Institute for Applied Language Studies, University of Edinburgh.

Ellis, N. (2005). At the interface: Dynamic interaction of explicit and implicit language knowledge. *Studies in Second Language Acquisition, 27,* 305-352.

Ellis, N. (2007). The associative-cognitive CREED. In B. Van Pattern & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 77-95). Mahwah, NJ: L. Erlbaum.

Ellis, R. (1991). Grammaticality judgments and second language acquisition. *Studies in Second Language Acquisition, 13*, 161-186.

Fujita, K., & Noro, T. (2009). The effects of 10-minute extensive reading on the reading speed, comprehension and motivation of Japanese high school EFL learners. *ARELE, 20,* 21-30.

Fukkink, R., Hulstijn, J., & Simis, A. (2005). Does training in second-language word recognition skills affect reading comprehension? An experimental study. *The Modern Language Journal, 89*, 54-75.

Grabe, W. & Stoller, F. (2011). *Teaching and researching reading*. Harlow: Longman.

Grabe, W. (2004). Research on teaching reading. *Annual Review of Applied Linguistics, 24*, 44-69

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Cambridge University Press.

Grabe, W. (2010). Fluency in reading – Thirty-five years later. *Reading in a Foreign Language 22,* 71-83.

Harrington, M. (2006). The lexical decision task as a measure of L2 proficiency. *EUROSLA Yearbook, 6*, 147-168.

Homan, S., Klesius, J., & Hite, C. (1993). Effects of repeated reading and nonrepetitive strategies on students' fluency and comprehension. *Journal of Educational Research, 87*, 94-99.

Hulstijn, J.H., Van Gelderen, A., & Schoonen, R. (2009). Automatization in second language acquisition: What does the coefficient of variation tell us? *Applied Psycholinguistics, 28,* 1-28.

Iwahori, Y. (2008). Developing reading fluency: A study of extensive reading in EFL. *Reading in a Foreign Language*, *20*, 70-91.

Kadota, S, Noro, T, Shiki, Mi. & Hase, N. (2014). *Vocabulary instruction on the basis of Computer-Based English Lexical Processing Test* (『英単語運用力判定テストを使った語彙指導』) Tokyo: Taishukan.

Kadota, S. (2010a). The interface between lexical and sentence processing in L2: An empirical study of Japanese EFL learners. *The Final Report Based on a Grant-in-Aid for Science Research (C) from 2007 to 2009*.

Kadota, S. (2010b). Reassessing the validity of the computer-based English lexical processing (CELP) test. *Studies in Foreign Languages and Cultures 15, (Published by School of Law and Politics, Kwansei Gakuin University), 91-106.*

Kame'enni , E. & Simmons, D. (2001) . Introduction to this special issue: The DNA of reading fluency. *Scientific Studies of Reading, 5*, 203-10.

Kanatani, K. (2008). *Zeal for English Education* (『英語教育熱 』). Tokyo: Kenkyuusha.

Klauda, S., & Guthrie, J. (2008). Relationships of three components of reading fluency to reading comprehension. *Journal of Educational Psychology, 100*, 310-321.

Kuhn, M., & Stahl, S. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology, 95*, 3-21.

Loewen, S. (2009). Grammaticality judgment tests and the measurement of implicit and explicit L2 knowledge. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philip, & H. Reinders, H. *Implicit and explicit in second language learning, testing and teaching* (pp. 94-112) . Bristol: Multilingual Matters.

Nakanishi, T. (2015). A meta-analysis of extensive reading research. *TESOL Quarterly*, *49*, 6-37.

Nassaji, H. (1998). *A component skills approach to adult ESL reading: Evidence from native speakers of Farsi*. Unpublished doctoral dissertation, Ontario Institute for Studies in Education, University of Toronto.

Nation, I.S.P. (2009). *Teaching ESL/EFL reading and writing.* New York: Routledge.

Nation, I.S.P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher,* 31 (7), 9-13.

National Reading Panel. (2000). *Teaching children to read: An evidence based assessment of the scientific research literature on reading and its implications for reading instruction* (National Institute of Health Pub. No. 004769). Washington, DC: National Institute of Child Health and Human Development.

Nuttall, C. (2005). *Teaching reading skills in a foreign language*. Oxford: Macmillan Education.

Perfetti, C.A. (1985). *Reading ability*. New York: Oxford University Press.

Perfetti, C.A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*, 357-383.

Perfetti, C.A. & Lesgold, A.M. (1977). Discourse comprehension and sources of individual differences. In M.A. Just & P.A. Carpenter (Eds.), *Cognitive processes in comprehension* (pp. 141-183). Hillsdale, NJ: Erlbaum.

Rasinski, T. (2003). *The fluent reader: Oral reading strategies for building word recognition, fluency, and comprehension*. New York: Scholastic Books.

Richards, J.C., & Schmidt, R. (2002). *Longman dictionary of language teaching & applied linguistics*. London, UK: Pearson Education.

Rumelhart, D.E. (1977). Toward an interactive model of reading. In S. Dornic & P. Rabbit (Eds.), *Attention and performance,* VI (pp. 573-603). Hillsdale, NJ: Erlbaum.

Samuels, S. (2002). Reading fluency: Its development and assessment. In S. Samuels & A. Farstrup (Eds.), *What research has to say about fluency instruction* (3rd ed., pp. 166-183). Newark, DE: International Reading Association.

Samuels, S.J.  (2006). Toward a model of reading fluency.  In S. Jay Samuels and Alan E. Farstrup (Eds.), *What research has to say about fluency instruction* (pp. 24-46). Newark, DE: International Reading Association.

Schmidt, R. (1992). Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition, 14*, 357-385.

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual.* Basingstoke: Palgrave Macmillan.

Segalowitz, N.S. (2003). Automaticity and attentional skill in fluent performance. In H. Liggenbach (Ed.), *Perspecctives on fluency* (pp. 200-219). Ann Arbor, MI: University of Michigan Press.

Segalowitz, N.S. (2003). Automaticity and second language. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 382-408). Oxford, UK: Blackwell.

Segalowitz, N.S., & Segalowitz, S.J. (1993). Skilled performance, practice, and the differentiation of speed-up from automatizatiion effects: Evidence from second language word recognition. *Applied Psycholinguistics, 14*, 369-385.

Shiotsu, T., & Weir, C.J. (2007). The relative significance of syntactic knowledge and vocabulary breadth in the prediction of reading comprehension test performance. *Language Testing*, *24*, 99-128.

Stanovich, K.E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency.  *Reading Research Quarterly*, *16*, 32-71.

Stanovich, K.E. (2000). *Progress in understanding reading*. New York: Guilford.

Sullivan, G. M. & Feinn, R. (2012). Using effect size—or why the *p* value is not enough. *Journal of Graduate Medical Education, 4*, 279–282.

Tono, Y. (2015).  *A new English vocabulary instruction to develop productive ability*  (『発信力をつける新しい英語語彙指導 』) . Tokyo: Sanseido.

Verhoeven, L. (2000). Components of early second language reading and spelling. *Scientific Studies of Reading, 4*, 313-330.