

## **Examining reading fluency in a foreign language: Effects of text segmentation on L2 readers**

Junko Yamashita  
Nagoya University  
Japan

Shingo Ichikawa  
Carnegie Mellon University  
United States

### **Abstract**

Grouping words into meaningful chunks is a fundamental process for fluent reading. The present study is an attempt to understand the relationship between chunking and second language (L2) reading fluency. The effects of text segmentation on comprehension, rate, and regression in L2 reading were investigated using a self-paced reading task in a moving-window condition. The participants were intermediate and advanced level Japanese EFL learners. The difficulty of chunking a text negatively affected comprehension and smoothness for the intermediate learners, while the advanced learners were able to overcome chunking difficulty. In this study, although the negative effects of chunking difficulty were observed, the positive effects of assisting chunking were not clearly detected, which was interpreted as suggesting that the relationship between chunking and reading needs to be considered in light of the complex interplay between text difficulty and different aspects of reading.

**Keywords:** reading fluency, chunking, L2 readers, comprehension, reading rate, backtracking, self-paced reading

Fluent readers read “smoothly, without hesitation and with comprehension” (Harris & Hodges, 1995, p. 85). Grouping words into grammatical and meaningful units or chunks is one of the fundamental processes to be automatized for fluent reading (Samuels, 2002). Although the importance of chunking ability for reading fluency is widely acknowledged, “research on the role of phrasing does not seem to be currently in vogue” (Rasinski, 2006, p. 16). The present study was motivated by the desire to understand second language (L2) learners’ fluency in reading.

### *Reading Fluency*

Reading fluency has recently become a hot topic in reading research and pedagogy (Samuels & Farstrup, 2006a). One of the driving forces behind this trend is a report by the National Reading

Panel in the US with the mission of proposing important factors for reading instruction (National Reading Panel, 2000). Reading fluency is recognized as one of the five essential elements to be considered in reading instruction, along with alphabets, comprehension, teacher education, and computer technology. Since then, reading fluency has provoked much discussion both theoretically and pedagogically (e.g., Breznitz, 2006; International Reading Association, 2002; Kuhn & Schwanenflugel, 2008; Kuhn & Stahl, 2004; Samuels, 2002; Samuels & Farstrup, 2006b; *Scientific Studies of Reading* (special edition), 2001; Rasinski, Blachowicz, & Lems, 2006).

In spite of this enthusiasm, reading fluency is a difficult concept to define. Researchers repeatedly mention the fluid nature of the concept (e.g., Breznitz, 2006; Kame'enui & Simmons, 2001; Samuels & Frastrup, 2006a; Schwanenflugel & Ruston, 2008; Wolf & Katzir-Cohen, 2001). This is partially because reading fluency is not a totally new concept but has appeared many times in the history of literacy education, and it has assumed different meanings according to social and educational needs or influential theories at different times (Rasinski, 2006). The different positions that researchers may take in their investigation seem another reason for the difficulty of achieving consensus in relation to the concept (Breznitz, 2006).

Despite this conceptual fluidity, a general consensus, at least concerning the end goal of fluency, seems to include both accuracy and speed factors (Wolf & Katzir-Cohen, 2001). Fluent readers are the readers who can comprehend a text's meaning smoothly and effortlessly at an appropriate rate. Reading slowly with a lot of halts and repetitions does not represent fluent reading even if the reader achieves higher levels of comprehension. Many researchers adopt the componential view of reading fluency: There are many components, or enabling skills, for achieving fluency in reading (e.g., Kuhn & Stahl, 2004; Pikulski, 2006; Schwanenflugel & Ruston, 2008; Topping, 2006; Wolf & Katzir-Cohen, 2001). There are, however, variations among researchers as to how many and what kinds of components they hypothesize. Kuhn and Stahl (2004) argue that the primary components which have achieved a certain consensus are "(a) accuracy in decoding, (b) automaticity in word recognition, and (c) the appropriate use of prosodic features such as stress, pitch, and appropriate text phrasing" (p. 416). The final goal of fluent reading (good text comprehension at an appropriate rate) is affected by the skills readers possess in each component. Moreover, influential components for reading fluency differ according to the developmental stage of readers (Wolf & Katzir-Cohen, 2001).

In the L2 reading field, despite some researchers' awareness of its importance for L2 pedagogy (e.g., Grabe, 2004, 2006), reading fluency has not yet received as much attention as it has in the field of first language (L1) reading. This is probably because L2 researchers and educators have tended to focus more on comprehension rather than rate and smoothness of reading, which is understandable in view of generally greater difficulty in L2 reading comprehension than in L1. However, understanding fluency in reading is no less important in L2 reading, because L2 learners desire not only to comprehend L2 texts well but also to read them quickly and smoothly. Finding out how comprehension and fluency interact with each other will help discover answers for several important questions for the further understanding of L2 reading; for example, whether L2 readers' comprehension, rate, and smoothness always go hand in hand or whether there is compensation among them and if so how different aspects of reading compensate for each other; and how or whether relationships among them change over the course of development. Although

the current study cannot firmly answer all these issues, it attempts to make a step towards a more comprehensive understanding of L2 reading by looking at different aspects of reading representing fluency. Among the three components identified by Kuhn and Stahl (2004), this study focuses on the third component (text phrasing). However, it does not deal with, at least explicitly, the use of prosodic clues. This is because, as will be discussed below, L2 readers' reliance on prosodic features is less clear compared with L1 readers.

### *Chunking and Reading*

Readers “who have not achieved fluency read either in a word-by-word manner or by grouping words in ways that deviate from the type of phrasing that occurs naturally in oral language” (Kuhn & Stahl, 2004, p. 418). Inappropriate skills in grouping words into units or chunks for reading are most easily manifested in oral reading (e.g., Amer, 1997; Schreiber, 1991). That is probably why the chunking ability is associated with prosodic reading in the L1 literature. For example, Allington (2006) states that “fluency is reading in phrases, with appropriate intonation and prosody—fluency is reading with expression” (p. 94). Prosody in the spoken language provides clues to syntactically and semantically meaningful chunks and children are particularly sensitive to prosodic clues in their sentence processing (Schreiber, 1987). When reading texts without overt prosodic clues, children have to find chunks in written sentences making use of different clues such as punctuation marks, morphosyntactic clues, or lexical information. Without this ability to chunk written texts, readers cannot read fluently even if they have mastered word decoding skills (Schreiber, 1991).

L2 learners may or may not rely on prosody as much as L1 children do because L2 learners do not necessarily acquire grammar and vocabulary in L2 spoken language before they start learning to read in L2. The relationship between the quality of oral reading and reading comprehension is less clear for L2 readers (Palumbo & Willcutt, 2006). Regardless of what knowledge sources (prosodic, morphosyntactic or lexical) readers may use, the ability to group words into appropriate units is an important determinant of fluent reading.

When investigating the relationship between chunking and reading, researchers have often manipulated the layout of a written text and observed the effects of such manipulation. The finding of the phenomenon that preorganization of texts into meaningful word groups can improve the reading behavior dates back several decades (Carver, 1970). However, very little is yet known about exactly how such manipulation affects reading. Answers to this question contain high educational value (e.g., how teachers can arrange written texts in order to help readers and how these different textual arrangements will help readers), and the literature is seeing increasing attempts to elucidate the effects of text segmentation on reading. This may be at least partly because the importance of reading fluency has been receiving larger amounts of attention, and chunking ability is an important subskill of fluent reading.

Studies along this renewed interest into the relationship between text segmentation and reading have not yet reached a firm consensus. However, we have seen that the effects of text segmentation vary for different readers and the effects appear over disparate aspects of reading. The text segmentation into meaningful chunks often helps readers whose skills are not sufficiently developed such as children (LeVasseur, Macaruso, Palumbo, & Shankweiler, 2006;

LeVasseur, Macaruso, & Shankweiler, 2008; O'Shea & Sindelar, 1983) and less skilled adult readers (Cromer, 1970). Skilled readers, on the contrary, tend to be resistant to any types of segmentation clues, both facilitative (Carver, 1970; Cromer, 1970) and disruptive (Cromer 1970). In other words, they can read well regardless of text conditions, probably because they have a well established chunking ability, and therefore either they do not need any more external clues or they can compensate for disruptive clues (segmentation of a text into meaningless word groups) using their own developed skills. Aspects of reading on which the effect of text segmentation has been reported include comprehension (Cromer 1970; O'Shea & Sindelar, 1983) and oral reading fluency such as fluency rating, false starts and word reading accuracy (LeVasseur et al., 2006; LeVasseur et al., 2008).

Regarding L2 reading, Kadota and his associates have reported a series of studies with Japanese learners of English. Kadota (1982) found that lower level readers achieved a higher level of comprehension when they read a text segmented into phrases with slashes (chunked text) than when they read an intact text and a text segmented into single words. Kadota and Tada (1992) presented texts under three conditions: sentence by sentence, phrase by phrase, and word by word. The mean recall score after reading was the highest under the phrase condition. In Kadota, Yoshida, and Yoshida (1999), texts were presented in three modes: Word by word, phrase by phrase, and clause by clause. Comprehension was significantly higher and reading time was significantly faster in the phrase and clause conditions than in the word condition. Hijikata (2005), however, who also examined Japanese university EFL learners, did not find a facilitative effect in chunking. She considered that the task may have been very challenging (the mean comprehension scores were below 50%) and readers could not utilize the chunking clues. The study suggests that the task difficulty, as well as readers' abilities, may influence the effect of chunking.

Among the studies reviewed above, the present study draws largely on Cromer (1970). Even though it may seem dated, there is important commonality with the present study; the basic research question (chunking ability and reading), the target population (adults), and the reading mode (silent reading). Some of the findings in the current study resembled those from Cromer. Therefore, his work is reviewed in detail below.

To test his hypothesis that skilled and less skilled readers differ in terms of how they internally organize reading materials, Cromer (1970) examined the effects of systematically encouraging or discouraging the use of meaningful chunks. The readers were college students who were divided into two pairs of skilled and less skilled readers according to their vocabulary and reading comprehension scores on a standardized test. As a result of this grouping there were four groups of readers called A, B, C, D, here for ease of interpretation. Group A consisted of skilled readers with high vocabulary scores, group B consisted of skilled readers with vocabulary scores less than group A, group C consisted of less skilled readers with vocabulary scores about the same as group B, and group D consisted of less skilled readers with the lowest vocabulary scores. Cromer presented texts in four modes: regular sentences, single words, meaningful phrases, and fragmented word groupings. That is, readers read a text sentence by sentence or word by word and so on. It was expected that the phrase mode would facilitate comprehension, but that the word mode and the fragment mode would interfere with comprehension because these modes violated the pattern necessary for adequate reading comprehension. Group A skilled readers

comprehended well in any mode (i.e., even in the word and fragment modes). These readers may be “able to adapt to these changes by imposing their own (effective) organization on the material” (Cromer, 1970, p. 481). However, group B skilled readers had lowered comprehension in the word and fragment conditions, displaying an expected pattern of responding to the imposed disruptive clues. Comprehension of group C less skilled readers was enhanced in the phrase mode to the extent that their level of comprehension even reached the same level as that of skilled readers. Their weak chunking skills were aided by external segmentation of texts.

Surprisingly, however, group D less skilled readers comprehended best in the word mode. Cromer speculated that these readers may have read each word more carefully under the word condition than they normally do. Furthermore, for this group of less skilled readers, grouping words into phrases or fragments had no significant effects in either facilitating or disrupting comprehension. It is possible, Cromer argued, that the readers always read in the same way as they did under the word mode (i.e., word by word) and how a text was presented, either in phrases or meaningless fragments, did not make any difference. This suggests that when readers do not possess sufficient vocabulary skills, chunking clues do not help. What they need first of all could be a sufficient word recognition skill.

The present study builds upon the previous studies that examined effects of text segmentation on reading, and expands the scope of investigation by including rate and smoothness as well as comprehension.

### *Reading Rate*

Our reading rate varies according to different purposes: for example, we skim newspaper headlines quickly or we prepare for exams by reading textbooks carefully and slowly. How can we accommodate this variation when we determine the general reading rate of each individual? Carver's (1990) distinction of five basic processes of reading—scanning, skimming, rauding, learning, and memorizing—provides a useful framework for understanding and measuring reading rate. These processes are ranked in a hierarchy in this order according to the involvement of more and more complex cognitive processes. For instance, scanning involves only lexical accessing, while skimming involves lexical accessing and semantic encoding. As a reflection of the increasing cognitive load, reading rates are hypothesized to vary in accordance with the hierarchy, with scanning the fastest (typically 600 wpm), and memorizing the slowest (typically 138 wpm). Variation in reading rates within an individual is basically regarded as the shift from one process to another.

Rauding, a term coined by Carver, in which readers read by looking at each consecutive word of a text and smoothly understand the complete thought contained in each sentence, is the most natural reading process used by many readers (at least by mature readers in L1). This process is used when we read texts well within our reading level. The typical rauding rate of adult L1 readers is about 300 wpm (Carver, 1990, p. 14), and the comprehension level is expected to be 75% or higher (pp. 143–144). With the aim of examining reading fluency, the current study attempts to elicit the rauding process of L2 readers, because it is the most typical reading process and we obviously read more slowly than usual in texts which are beyond our reading ability. Efforts to make reading texts match L2 readers' ability seem important, because there is no

simple formula to estimate grade-equivalent text difficulties for L2 readers. Sometimes very slow L2 reading rates are reported (e.g., 64 to 85 wpm: L1 Japanese university students reading English) (Taguchi, Takasu-Maass & Gorsuch, 2004), but the comprehension level is also low (below 30%).

Having said that, however, reading rates reported in studies of L2 reading are, in general, much slower than a L1 native speaker norm even if L2 readers are given simple texts—sometimes even without comprehension questions. It is often the case that the reading rate does not reach 200 wpm (e.g., 183 wpm for L1 Chinese university students reading English, Fraser, 2007), or is even below 100 wpm, such as 84 to 113 wpm in L1 Japanese secondary school students reading English (Iwahori, 2008). Exceptions may include proficient bilingual populations whose L2 is linguistically closer to L1 such as English and French, although such a population may still read slowly compared to native speakers of a language (see summary in Fraser, 2007). It is difficult to specify factors that are responsible for the large gap between L1 and L2 reading rates, but they would include variables such as L2 proficiency, L1-L2 distance, and L2 learning environment. Variation among the studies can also be attributed to reading tasks or materials employed in different studies.

Since readers shift their reading processes according to the purposes, an ideal condition to measure the reading rate may be not giving comprehension questions (Carver, 1990; Fraser, 2007). However, some simple comprehension measures can still be used, for instance, “relatively easy yes-no questions after each passage that were general in scope” (Carver, 1990, p. 192). The present study utilized such simple comprehension questions because we included comprehension as one of the dependent variables, but efforts were made to make the questions as simple as suggested above.

### *The Study*

As mentioned above, the purpose of the study is to expand our understanding of reading fluency by focusing on chunking ability. Adopting the text segmentation method, four text modes are used: the Whole-Text, Single Words (word-by-word), Meaningful Chunks (chunk-by-chunk), and Fragmented Word-groups (fragment-by-fragment). The Whole-text mode most closely represents a natural reading condition. The single word mode is supposed to force readers to read word by word. As former studies have shown, this mode is expected to be a challenging condition for readers who have acquired skills to read in larger meaningful units. However, if readers always read word by word as suggested by the weak readers in Cromer (1970), this mode may not be so challenging. The Chunk mode simulates the case when learners are provided with assistance in chunking. The Fragment mode is an unnatural condition, but it simulates the case when learners segment texts into inappropriate word-groups, which is often observed in learners' oral reading with inappropriate phrasing (e.g., Amer, 1997). This kind of disruption of phrase boundaries also occurs frequently at the end of lines in printed texts (LeVasseur et al., 2006). Since reading was the target process, the study attempted to use relatively simple texts from the linguistic and cognitive points of view so that the difficulty levels of texts were within L2 readers' reading ability.

We use the word “chunk,” rather than “phrase,” to refer to a meaningful multi-word unit in a

written text, because the term chunk seems to accommodate a wider range of concepts concerning multi-word units being discussed in recent studies (e.g., Schmitt, 2004). Although chunks in the present study largely overlap phrases, the intention is that chunks will be viewed as not necessarily syntax-bound and will include a wider range of concepts.

In order to help guide the study, the results will be examined according to the following hypotheses:

1. Appropriately chunked texts facilitate reading at a lower level of L2 proficiency.
2. Appropriately chunked texts do not facilitate reading at a higher level of L2 proficiency.
3. Inappropriately chunked texts interfere with reading at a lower level of L2 proficiency.
4. Inappropriately chunked texts do not interfere with reading at a higher level of L2 proficiency.

These hypotheses are based on the assumption that learners at an advanced level have higher levels of chunking skills available for utilization in their reading than learners at a lower L2 proficiency level. However, it must be noted that, since the difficulty level of texts is made to be within readers' ability and comprehension questions only require general or literal understanding, hypotheses 1 and 2 may not be evaluated by using the comprehension scores. This is because, if readers have achieved high levels of comprehension, it would be difficult to observe improvement in their comprehension scores. Hypotheses 3 and 4, on the other hand, can still be valid hypotheses, because it is possible to decrease their level of comprehension if inappropriate chunking clues hinder their comprehension. There is no such concern with regard to reading rate and regression, because we can still expect both positive (e.g., achieving the same level of comprehension more quickly and smoothly) and negative (e.g., achieving the same level of comprehension more slowly and less smoothly) effects on these two indices.

## Method

### *Participants*

Forty-eight Japanese university students participated in the study. A 58-item cloze test was prepared to measure their overall L2 proficiency (mean = 19.81, *SD* = 8.92, Chronbach's alpha = 0.91). The test items were selected from the Placement/progress test (Test B) originally constructed in the Edinburgh Project on Extensive Reading (EPER) for the purpose of measuring a complete range of English language proficiency (EPER, not dated). The original test consists of 147 items based on 13 relatively short passages. In the past years, the first author had given this test to Japanese students who were similar to the present participants. On the basis of the data collected in the past and using reliability analysis of the SPSS statistical package, items contributing to the test's internal consistency were selected, which resulted in the 58-item cloze test based on five passages. The motivation of selectively shortening the original test was to reduce the imposition of test taking on the participants whilst measuring their proficiency as reliably as possible. This shortened test was used in a previous study of the first author, where it

demonstrated high internal consistency and differentiated learners at different levels of proficiency well (Yamashita, 2007). Therefore, the test was considered appropriate to be adopted in the present study. The participants were divided into two levels by the mean score (intermediate and advanced,  $n = 24$  in each). The test results and characteristics of the two groups, which were obtained by a questionnaire, are summarized in Table 1.

Table 1. *Characteristics of participants*

	Group	Min	Max	Mean	SD
Test result	Advanced	20	44	26.92	6.94
	Intermediate	6	18	12.71	3.04
Age	Advanced	18	35	25.63	4.62
	Intermediate	18	24	19.63	1.88
Years of English study	Advanced	6	22	12.67	3.89
	Intermediate	6	11	7.00	1.44

*Note.*  $n = 24$  for each group

### *Materials*

Reading materials were chosen from a standardized test of English proficiency widely used in Japan (The Eiken Test in Practical English Proficiency). The test is constructed by the Society for Testing English Proficiency (STEP), Japan's largest testing organization. The test aims at testing four skills as well as grammar and vocabulary. Currently seven bands are distinguished, with each corresponding to a different school grade in Japan. Except for the test of speaking ability, it is a multiple-choice test. It is administered three times a year all over the country, and the items are disclosed to the public soon after it is finished.

The pre-first grade is said to correspond to the intermediate level of university students. The initial passages for the present study were selected from this grade. The selected passages were originally for the listening comprehension section, but these were judged to better serve the present study than the ones for the reading comprehension section for the following reasons. First, the passages were linguistically less complex, and seemed more appropriate to investigate reading fluency for the reason mentioned above. Second, the length of the texts was shorter and seemed more appropriate than the passages in the reading section, which were too lengthy and would have made the experimental task too tiring.

Descriptive and narrative type texts were chosen and pilot-tested with several groups of university students who were similar to those involved in the main study. The learners in the pilot study were asked to read the passages and estimate topic familiarity and level of difficulty (each on a 7-point Likert scale), and indicate unknown words by circling them. Comparing results of several passages, four texts were finally selected. Words which were indicated as unknown by more than one-fourth of the final group of 42 students in the pilot study were replaced by more frequent synonyms, or a part of the text was rewritten to avoid using the indicated words. Table 2 lists the characteristics of the four passages after such modifications were made. Means and standard deviations for the topic familiarity and perceived level of difficulty are based on the responses from the final 42 students. One-way ANOVAs indicated there was no significant difference among the four texts in terms of familiarity of the topic and



the perceived level of difficulty.

Table 2. *Characteristics of reading materials*

	Text 1	Text 2	Text 3	Text 4
Number of words	131	137	134	131
Number of letters	652	698	679	663
Flesch reading ease	47.4	46.5	45.5	44.9
Flesch-Kincaid grade level	10.2	10.9	11	11
Topic familiarity	2.10 (1.28)	2.33 (1.18)	1.95 (1.17)	1.81 (1.04)
Perceived difficulty level	4.26 (1.21)	3.83 (1.19)	4.23 (1.3)	3.81 (1.06)

*Note.* Values presented for Topic familiarity and Perceived difficulty level represent means on the following scales with standard deviations enclosed in parentheses.

Topic familiarity: 1 = not familiar at all, 7 = very familiar.

Perceived difficulty level: 1 = very difficult, 7 = very easy.

All these texts were presented in the four modes (Whole, Word, Chunk, and Fragment). Examples of chunked and fragmented texts are shown below. Chunk boundaries mostly corresponded with phrase boundaries (see Appendix), and fragments deliberately violated such grammatical units. The number of chunks and fragments in one sentence was made identical. Therefore, there were equal numbers of chunks and fragments in one text, which varied from 30 to 35 across the four texts.

Chunked:

The origin of Australian Rules Football / is unclear. / Some people say / it might have developed / from an ancient game / in which a ball made of kangaroo skin / was kicked around.

Fragmented:

The origin of Australian / Rules Football is unclear. / Some people say it might / have developed from an / ancient game in / which a ball made of kangaroo / skin was kicked around.

Five true-false questions were created for each text. (The original test had only two multiple-choice questions for each text, and this was felt insufficient.) Questions were made so that readers would be able to answer correctly as long as they understood the literal meaning of each passage. Efforts were also made to assure that the questions could not be answered without reading the texts. The questions were printed on a sheet of paper in Japanese to be given to readers after they had finished the reading task on a computer.

### *Procedure*

The task was a self-paced reading under a moving-window condition (Just, Carpenter, & Woolley, 1982). The program developed for the present study was run on a personal computer. The data were taken individually in a quiet room by the first author. A single session lasted about 30 minutes, including a demographic questionnaire, explanation, practice, the main task, and a debriefing interview. All participants were paid for their participation.

In the practice session, each participant was trained using practice passages. All participants easily understood the task and quickly mastered how to read forward and backward on the computer by pressing keys. No technical difficulties were reported in the interview. In the main study session, each participant read a different text under each of the four conditions. The Whole-Text condition—the most natural mode—was always presented first, and the order of the other three conditions was counter-balanced. The title of each text was given to the readers with its Japanese translation before they started reading, so that they, particularly weaker students, could all have a similar cognitive framework for the texts.

The participants were instructed to read as quickly as possible, but to make sure they understood the texts. They were told that there would be comprehension questions after reading, that the questions were simple and they would be able to answer without too much difficulty as long as they understood the general meaning of the text. Only the title of a text appeared on the first screen that the readers faced. In the Whole condition, the entire text appeared on the screen when readers pressed the right arrow button on the keyboard. The readers did not press any key while reading, and finally pressed the right arrow key again when they had finished reading. In the other three conditions (Single Word, Chunk, and Fragment), the readers only saw a series of underlines representing the position of words on the screen after they pressed the first button. When the readers pressed the right arrow key a second time, the first presentation unit appeared (a word, a chunk, or a fragment) replacing the underline(s) corresponding to that unit. When the readers pressed the key to request the next unit, the next unit appeared and the previous unit was replaced by its underline(s). Thus only one reading unit was visible on the screen at any time. If the readers wanted to re-read, they requested the previous unit by pressing the left arrow key, then the previous unit came back to the screen, replacing the underline(s). Readers were able to go back to read as many units as they liked. The readers kept reading texts this way by displaying one unit after another at their own pace.

After reading a text, the readers answered true-false questions without referring to the passage. The computer program recorded the time between the two keystrokes (the times recorded were added in the end to obtain the entire time to finish reading a passage) and the direction of reading (right and left arrow keystrokes). The level of reading comprehension was measured by the accuracy rate on the true-false questions.

A potential problem of this key pressing method and how it affected the way we examined reading rate results should be discussed. The numbers of keystrokes while learners were reading the texts, which approximates the numbers of presentation units, were different across the four modes (zero in the Whole-Text mode, about 130 times in the Single Word mode, and about 30 times in the Chunk and Fragment modes), although the numbers in Chunk and Fragment modes were identical because there were equal numbers of chunks and fragments. We were not able to eliminate the time spent on key pressing. As will be reported in the result section, both groups of readers were slowest in the Single Word mode. It could be argued that the large number of keystrokes in this mode might have added extra time, and possibly that this reading condition forcing the readers to keep pressing the key throughout their reading might have disrupted their normal reading behavior. In order to address this possibility, the hypotheses were examined based on the results of Chunk and Fragment modes when reading rate was considered. However,

we can still compare the result of the Single Word mode between the groups, because the condition, including these potential problems, equally applies to each group.

## Results

Table 3 lists the means of comprehension scores, reading rates (words per minutes), and regression indices (the number of words the readers backtracked by pressing the left arrow key) for each group under all conditions. Before the main analysis, the possibility of text specific effect was tested. One-way ANOVAs were applied to examine the difference in the comprehension score, rate, and regression according to differences among texts. No significant effect related to the differing texts was observed for any of the three dependent variables. This result was the same when the data were analyzed for the intermediate and advanced groups separately. Therefore, we can say that differences observed in the following analyses are not due to differences in the individual texts per se, but most likely due to the independent variables—presentation mode and levels of L2 proficiency.

Table 3. Mean comprehension scores, reading rates, and regression indexes for each group under all conditions

Modes	Comprehension		Rate		Regression	
	Advanced	Intermediate	Advanced	Intermediate	Advanced	Intermediate
Whole	4.25 (0.94)	4.33 (0.82)	133 (37.49)	85 (19.93)	---	---
Word	4.38 (0.92)	4.50 (0.59)	105 (28.45)	71 (20.50)	6.83 (20.33)	20.25 (33.22)
Chunk	4.33 (0.70)	4.29 (0.69)	131 (35.68)	79 (28.18)	2.67 (6.60)	31.58 (40.19)
Fragment	4.38 (0.82)	3.58 (1.25)	125 (39.64)	76 (23.31)	9.00 (20.14)	40.50 (59.81)

*Note.* Dashes indicate regression values are not available under the Whole condition. Values enclosed in parentheses represent standard deviations. Reading rate is represented in wpm. Regression is presented as the number of words regressed.

### Comprehension

The mean comprehension scores were generally high (more than 80%), and even the lowest score maintained over 70% accuracy. It could be said that the texts were within the participants' reading comprehension ability, regardless of their L2 proficiency, at least in terms of understanding their literal meanings.

A 4 x 2 ANOVA was computed, with the presentation mode as the within-subjects variable and the learner groups as the between-subjects variable. There was no significant main effect of Mode and of Group,  $F(3, 138) = 2.633, p = .052, \eta_p^2 = .054$ ;  $F(1, 46) = 1.297, p = .261, \eta_p^2 = .027$ , but the interaction between Mode and Group was significant,  $F(3, 138) = 3.194, p < .05, \eta_p^2 = .065$ . The effect sizes of Mode, Group, and their interaction were all small. This was

probably because the comprehension scores were generally high and there was not much variability in the data. Post hoc analysis (Bonferroni) indicated that the comprehension score of the intermediate group was significantly lower than that of the advanced group under the Fragment mode ( $p < .05$ ). Otherwise, there was no difference between the groups. Also the comprehension score of the intermediate group was significantly lower in the Fragment mode than in the other modes ( $p < .05$ ). In other words, although the accuracy level of 71.6% (3.86) may be reasonably high, the intermediate group lowered their comprehension in Fragment mode. There was no such difference across the modes for the advanced group.

### *Rate*

Considering generally high comprehension scores, it could be said that the reading rates recorded here reflect these L2 readers' reading rate when they are reading easy texts for general comprehension. A 4 x 2 ANOVA was used with the presentation mode and the groups as independent variables. There was a main effect of Mode,  $F(3, 138) = 25.567, p < .001, \eta_p^2 = .357$ , and of Group,  $F(1, 46) = 32.300, p < .001, \eta_p^2 = .413$ . The interaction between Mode and Group was significant,  $F(3, 138) = 5.041, p < .01, \eta_p^2 = .099$ . The effect sizes show that the main effects of Mode and Group were indeed large, but that of the interaction was relatively small. Post hoc analysis (Bonferroni) revealed that the advanced group read significantly faster than the intermediate group in all four modes ( $p < .001$ ). There was no difference among Whole-Text, Chunk and Fragment modes for either group. The advanced group was significantly slower in the Single Word mode than all the other modes ( $p < .001$ ). Although the intermediate group was also the slowest in the Single Word mode, the difference was significant only in comparison with the Whole mode ( $p < .01$ ). There was no significant difference among the Single Word, Chunk and Fragment modes for this group. The result suggests that the significant interaction between Mode and Group can be largely attributed to the advanced group slowing down substantially in the Single Word mode in comparison with the intermediate group. On the other hand, the intermediate group was relatively constant in their rate across modes.

Principal findings are summarized as follows: (1) The advanced group consistently read faster than the intermediate group regardless of the conditions, (2) reading rates were not different between the Chunk and Fragment modes for both groups, (3) the extent to which the readers slowed down in the Single Word mode in comparison with other modes was greater for the advanced group than for the intermediate group.

### *Regression*

While not all readers in the two groups regressed in their reading, in general, a larger number of students backtracked in the intermediate group than in the advanced group. The number of students who regressed in the Single Word, Chunk, and Fragment modes respectively was 10, 6, and 6 in the advanced group and 16, 17, and 17 in the intermediate group.

Table 3 shows large standard deviations, particularly for intermediate learners. There were considerable individual differences. For example, the largest range varied from 0 to 98 (advanced group in Single Word mode) and 0 to 202 (intermediate group in Fragment mode). Such individual variation seems to have resulted in the large deviation from the mean.

A 3 x 2 ANOVA with the three modes (Single Word, Chunk, and Fragment) as the within-subjects variable and the two groups as the between-subjects variable was computed. There was a significant main effect of Mode,  $F(2, 92) = 3.140, p < .05, \eta_p^2 = .064$ , and of Group,  $F(1, 46) = 8.461, p < .01, \eta_p^2 = .155$ . The interaction between Mode and Group was not significant,  $F(2, 92) = 2.291, p = .107, \eta_p^2 = .047$ . Effect sizes suggest that the effect of Mode, although statistically significant, was not so large as that of Group. Post-hoc analysis (Bonferroni) indicated that there was no difference among the three modes for the advanced group. The intermediate group backtracked significantly more in the Fragment mode compared with the Single Word mode ( $p < .05$ ), but there was no significant difference between the Word and Chunk modes, or between Chunk and Fragment modes. The difference between the groups was significant under the Chunk and Fragment modes ( $p < .01, p < .05$ ). There was no significant difference between the groups in the Single Word mode.

Principal findings can be summarized as follows: (1) The advanced group generally backtracked to a lesser extent than the intermediate group, although there was no statistically significant difference between them in the Single Word mode, and (2) as will be discussed below, the backtracking behavior of the advanced group was not affected by different chunking clues, while that of the intermediate group was.

## Discussion

The study examined how chunking clues affect comprehension, rate, and regression of L2 readers at different levels of L2 proficiency. We will first discuss three aspects of reading separately according to the four hypotheses, then integrate the results and discuss the reading behaviors of the L2 readers.

Regarding comprehension, testing hypotheses 1 and 2—effects of appropriate chunking clues on lower and higher level learners—was not possible. As explained above, this was due to the fact that the texts were simple and comprehension scores reached a ceiling. On the other hand, the negative effect of the Fragment mode on comprehension was evident in the intermediate group. Hypothesis 3—negative effect of inappropriate clues on lower level learners—was supported, indicating that comprehension was disrupted when making appropriate chunks in a text was difficult. In contrast, there was no such negative effect on the advanced learners and hypothesis 4—no effect of inappropriate clues on higher level learners—was supported. We could say that the advanced learners must have possessed a superior chunking ability and were able to compensate for the inappropriate text segmentation successfully.

The reading rate result did not support hypotheses 1 and 3—positive and negative effects of appropriate and inappropriate clues on lower level learners—since the rates of the intermediate level learners in the Chunk and Fragment modes did not significantly differ. On the other hand, hypotheses 2 and 4—no positive and negative effects of appropriate and inappropriate clues on higher level learners—were supported, because the reading rates of the advanced group in the Chunk and Fragment modes did not show a significant difference either. Although the variation of reading rate across different modes was similar between the groups, it must be stressed that

the advanced group read constantly faster than the intermediate group. Any change in the modes did not affect their absolute advantage over the intermediate group. Thus, the advanced group must have possessed skills to read English texts faster than the intermediate group.

In addition to the possible disruption of continued key pressing, the slowest rate in the Single Word mode for both groups may be explained by the concept of limited perceptual span<sup>1</sup>. Using the moving window technique, McConkie and Rayner (1975) found that L1 readers reduced their reading rate by about 60% when they could see only 17 character spaces (i.e., about two to three words) at a time. Compared with their result, the magnitude of decrease was smaller in this study—21% and 16% for the advanced group and the intermediate group respectively. However, the phenomenon that the limited perceptual span reduces the reading rate seems to apply to both groups of L2 readers.

In spite of the difficulty of the limited perceptual span, the advanced group was still significantly faster than the intermediate group. The advanced learners must have been able to continue reading a single word, one after another, quickly and accurately. Therefore, we can say that they were probably more efficient than the intermediate group in word recognition skills.

An important finding indicative of the difference in the reading process between the groups is the fact that the advanced learners reduced their rate in the Single Word mode to a larger extent than the intermediate group. This is probably because, since the advanced group possessed more efficient word recognition skills and they could normally read in a larger unit, they were more disrupted by the forced narrowness of the window. The intermediate group, on the other hand, was less sensitive to the narrowed perceptual span. When word recognition skills are less efficient and automatized to a lesser extent, readers use more of their cognitive resources to read each word. This not only reduces reading rate but also takes their attention away from the larger context. The result that there was no significant difference in rate among Single Word, Chunk, and Fragment modes for the intermediate group endorses their insensitivity to the larger context. This finding is relevant to more fundamental components of reading fluency—word level processing—rather than directly to chunking, and adds an interesting discovery to the study.

Regarding regression, the statistical test supported hypotheses 2 and 4—no positive and negative effects of appropriate and inappropriate clues on higher level learners—since the advanced group did not show any difference in the amount of regression among the Word, Chunk, and Fragment modes. Hypotheses 1 and 3—positive and negative effects of appropriate and inappropriate clues on lower level learners—were also supported as discussed below, but further consideration is necessary.

Theoretically speaking, the Single Word and Fragment modes are inappropriately chunked conditions, since reading units in both modes deviate from those normally required for adequate reading comprehension. The Fragment mode is supposed to be more demanding of the two because it deliberately discourages appropriate chunking. The regression behavior of the advanced group showed a tendency that supports this assumption: They backtracked least in the Chunk mode, more in the Single Word mode, and most in the Fragment mode. However, this was not the case for the intermediate group: they backtracked least in the Single Word mode, more in the Chunk mode, and most in the Fragment mode. The present study cannot answer why

the intermediate group had a tendency to backtrack more in the Chunk mode than in the Single Word mode. However, as the result of rate suggests, this group's less efficient word processing skills might provide a partial answer. Reading multiple words as a unit could be more demanding for such readers than reading word by word, possibly because their less efficient word recognition processes use up their cognitive resources and they do not have enough attention left to integrate words into a coherent meaning representation. As reviewed earlier, a group of less skilled readers in Cromer (1970) achieved their highest level of comprehension when they read texts word for word in comparison with other conditions such as sentences, phrases, and fragments. However, there may be a stage in the course of development when focusing the attention on single words helps readers. Although whether this is true remains a question to be explored in the future, the backtracking behavior of the intermediate group suggests that they may have felt the least difficulty in the Word mode. Therefore, hypotheses 1 and 3 are examined based on the expected positive and negative effects of Chunk and Fragment modes by regarding the Word modes as a base line.

The intermediate group did not regress significantly more in the Chunk mode than in the Single Word mode, but they did in the Fragment mode. Although the positive effect of Chunk mode is rather weak because there was not a significant difference between the Chunk and Fragment modes, the backtracking behavior of the intermediate learners can be said to be sensitive to the difference between the Chunk and Fragment modes. Thus hypothesis 3 was supported by the negative effect of the Fragment mode, and hypothesis 1 was supported indirectly, not by the facilitative effect in the Chunk mode as originally expected, but by the lack of negative effect in the Chunk mode compared with the Fragment mode.

The overall results of the intermediate group indicate that disruption in constructing meaningful chunks in the reading process negatively influences comprehension and smoothness of reading. Although the Fragment mode was experimentally created and learners would not read texts written in that manner in natural situations, as mentioned above, this mode was meant to simulate the situation when learners wrongly segment texts either because their chunking ability is weak or because the texts are beyond their reading ability. In such cases, texts might look like a meaningless string of words or word-groups to readers. Unlike these relatively straightforward results in comprehension and regression, it is surprising that the intermediate group did not slow down in the Fragment mode in spite of their difficulty in comprehension and larger amount of backtracking. Although this remains a question to be answered, part of the reason may be that the texts used were still relatively easy and the comprehension level maintained at over 70%.

Regarding facilitative effects of Chunk mode on the intermediate group, the study observed only indirect evidence in terms of regression. There was no facilitation on reading rate. There are at least two potential reasons for this. First, they have less efficient word recognition skills. Although the present study focused on chunking ability, when more fundamental word-level skills are insufficiently developed, readers may not be able to benefit from clues targeting at the higher level processing. Secondly, the change of reading rate may need to be considered in relation to comprehension: Only when chunking clues help comprehension, may the reading rate improve.

A series of studies on L1 children's oral reading fluency have suggested that chunking ability

may not relate to reading comprehension when texts are simple and straightforward (Schwanenflugel, Hamilton, Kuhn, Wisenbaker, & Stahl, 2004), but that it can become an indicator of comprehension when children read more complex ones (Miller & Schwanenflugel, 2006; Schwanenflugel, Meisinger, Wisenbaker, Kuhl, Strauss, & Morris, 2006). These L1 studies and our speculation described above about the interplay between the rate change and comprehension difficulty point to the necessity of examining the relationship among chunking clues, comprehension, and rate using more difficult texts.

Compared with the intermediate group, the advanced group achieved superior performance (consistently high comprehension, faster rate, and generally less regression). This could partly be attributed to their higher chunking skills, which was reflected in their ability to overcome text presentation distortion. In addition, they may also be superior at the word-level processing.

The present study attempted to expand our understanding of the relationship between chunking and L2 reading fluency. Although there are several limitations as discussed below, the study has increased our current knowledge of L2 reading by demonstrating the following points. First, even if readers achieve the same level of comprehension by reading relatively simple texts, the rate and smoothness of reading are considerably different at different stages of development. Second, appropriate chunking of a text is indeed crucial for L2 reading. This then supports one of the components of reading fluency proposed by Kuhn and Stahl (2004). Third, limiting the perceptual span makes reading difficult, and more advanced level learners suffer more from this difficulty. This is suggestive of the development in reading skill, that is, the advanced learners have more efficient word processing skills and therefore can integrate larger numbers of words when reading. Fourth, the difficulty of chunking a text does not necessarily affect different aspects of reading in the same way, which indicates the complexity of the interaction among comprehension, rate and smoothness of reading. This then calls upon the necessity of looking at multiple facets of reading when we try to understand reading fluency.

Finally, we discuss limitations of the present study, which should be addressed in the future. Firstly, as mentioned earlier, the change of reading rate may need to be examined in the interaction with comprehension using more complex and demanding texts. Secondly, the current data suggested that the intermediate learners were less efficient in word recognition skills. The design of the study, however, does not allow us to conclude definitely that this is the case. Although the study targeted the text phrasing, more fundamental word level skills will need to be included in order to fully explain why lower level readers are less fluent. Thirdly, in relation to the second point, chunking ability was not measured but instead inferred from general L2 proficiency and from the processing skills of appropriately and inappropriately chunked texts. Measuring chunking ability would be a more direct way of dealing with the relationship between chunking and reading.<sup>2</sup> One of the difficulties of this approach may be that currently we do not clearly know how to measure chunking ability per se. Probably it overlaps with language proficiency and reading ability, as suggested in Cromer (1970) and the present study. Many recent L1 studies are interested in children's oral reading fluency, and chunking ability seems included within the measure of prosodic reading. Whether L2 researchers follow this line of investigation or they use a different method in the course of silent reading is something that L2 studies have to consider carefully. Fourthly, there was a problem in the key-pressing method. Forced repetitive key-pressing might have been disruptive for readers. Less intruding methods



such as eye-movement tracking need to be considered for further studies. Finally, regression during reading is normally examined by eye-movements (e.g., McConkie & Rayner, 1975; Rayner, Chace, Slattery, & Ashby, 2006; Rayner & Sereno, 1994). Measuring it in the way in which this study did is a very new approach. Although some reasonable behavioral results following the expected difficulty of reading conditions were observed, the validity of the current method is not established. Therefore it is fair to raise a caution. Efforts should be made to find out the validity of this method.

## Conclusion

Reading in meaningful chunks is one of the fundamental processes necessary for fluent reading. The present study is an attempt to understand the relationship between chunking and reading fluency of L2 readers. Reading fluency requires us to see reading multidimensionally. It thus can provide a framework for a more comprehensive understanding of L2 reading, which would help learners achieve smooth comprehension in reading.

## Acknowledgments

This research was supported by a Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (No. 16520339). The authors would like to thank Masatoshi Sugiura and Shogo Shirai for the program used for data elicitation. The authors also appreciate constructive comments from anonymous reviewers of *Reading in a Foreign Language*.

## Notes

1. The authors owe this idea to one of the anonymous reviewers.
2. This idea was suggested by one of the anonymous reviewers.

## References

- Amer, A. A. (1997). The effect of the teacher's reading aloud on the reading comprehension of EFL students. *ELT Journal*, 51, 43–47.
- Allington, R. L. (2006). Fluency: Still waiting after all these years. In S. J. Samuels & A. E. Farstrup (Eds.), *What research has to say about fluency instruction*. (pp. 94–105). Newark, DE: International Reading Association.
- Breznitz, A. (2006). *Fluency in reading: Synchronization of processes*. Mahwah, NJ: Lawrence Erlbaum.
- Carver, C. (1970). Effect of a “chunked” typography on reading rate and comprehension. *Journal of Applied Psychology*, 54, 288–296.
- Carver, C. (1990). *Reading rate: A review of research and theory*. New York: Academic Press.
- Cromer, W. (1970). The difference model: A new explanation for some reading difficulties. *Journal of Educational Psychology*, 61, 471–483.
- Edinburgh Project on Extensive Reading (not dated). *EPER publications: Supporting English*

- reading schemes for schools and colleges world-wide* [Brochure]. Edinburgh: University of Edinburgh.
- Fraser, C. A. (2007). Reading rate in L1 Mandarin Chinese and L2 English across five reading tasks. *The Modern Language Journal*, 91, 372–394.
- Grabe, W. (2004). Research on teaching reading. *Annual Review of Applied Linguistics*, 24, 44–69.
- Grabe, W. (2006). An agenda for reading instruction research. Lecture given at The 2006 Asia TEFL International Conference, Fukuoka, Japan.
- Harris, T. L., & Hodges, R. E. (1995). *The literacy dictionary*. Newark, DE: International Reading Association.
- Hijikata, Y. (2005). The chunking process and reading comprehension of Japanese EFL learners. *Annual Review of English Language Education in Japan*, 16, 61–70.
- International Reading Association. (2002). *Evidence-based reading instruction: Putting the national reading panel report into practice*. Newark, DE: International Reading Association.
- Iwahori, Y. (2008). Developing reading fluency: A study of extensive reading in EFL. *Reading in a Foreign Language*, 20, 70–91.
- Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, 111, 228–238.
- Kadota, S. (1982). Some psycholinguistic experiments on the process of reading comprehension. *Journal of Assumption Junior College*, 9, 49–70.
- Kadota, S., & Tada, M. (1992). Eibun oyobi nihonbun no dokkai to shoritani [Reading comprehension and processing units in English and Japanese]. *Annual Bulletin of Research Institute for Social Science*, 22, 137–153.
- Kadota, S., Yoshida, S., & Yoshida, H. (1999). Dokkai niokeru shoritani—Eibun no teijitani garikaido oyobi shorijikan ni oyobosu eikyo [Processing units in EFL reading: An effect of presentation units on comprehension rate and time]. *Annual Review of English Language Education in Japan*, 10, 61–71.
- Kame'enui, E., & Simmons, D. C. (2001). Introduction to this special issue: The DNA of reading fluency. *Scientific Studies of Reading*, 5, 203–210.
- Kuhn, M. R., & Stahl, S. A. (2004). Fluency: A review of developmental and remedial practices. In R. B. Ruddell & N. J. Unrau (Eds.), *Theoretical models and processes of reading* (5th ed.) (pp. 412–453). Newark, DE: International Reading Association.
- Kuhn, M. R., & Schwanenflugel, P. J. (Eds.). (2008). *Fluency in the classroom*. New York: The Guilford Press.
- Le Vasseur, V. M., Macaruso, P., Palumbo, L. C., & Shankweiler, D. (2006). Syntactically cued text facilitates oral reading fluency in developing readers. *Applied Psycholinguistics*, 27, 423–445.
- LeVasseur, V. M., Macaruso, P., & Shankweiler, D. (2008). Promoting gains in reading fluency: A comparison of three approaches. *Reading and Writing: An Interdisciplinary Journal*, 21, 205–230.
- McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics*, 17, 578–586.
- Miller, J., & Schwanenflugel, P. J. (2006). Prosody of syntactically complex sentences in the oral reading of young children. *Journal of Educational Psychology*, 98, 839–843.
- National Reading Panel. (2000). *Teaching children to read: Reports of the subgroups* (No. 00-

- 4754). Washington, DC: National Institute of Child Health and Human Development. Retrieved from: [http://www.nichd.nih.gov/publications/nrp/upload/report\\_pdf.pdf](http://www.nichd.nih.gov/publications/nrp/upload/report_pdf.pdf)
- O'Shea, L. J., & Sindelar, P. T. (1983). The effects of segmenting written discourse on the reading comprehension of low- and high-performance readers. *Reading Research Quarterly, 18*, 458–465.
- Palumbo, T. J., & Willcutt, J. R. (2006). Perspectives on fluency: English-language learners and students with dyslexia. In S.J. Samuels & E. Farstrup (Eds.), *What research has to say about fluency instruction* (pp. 159–178). Newark, DE: International Reading Association.
- Pikulski, J. J. (2006). Fluency: A developmental and language perspective. In S. J. Samuels & E. Farstrup (Eds.), *What research has to say about fluency instruction* (pp. 70–93). Newark, DE: International Reading Association.
- Rasinski, R. V. (2006). A brief history of reading fluency. In S. J. Samuels & E. Farstrup (Eds.), *What research has to say about fluency instruction* (pp. 4–23). Newark, DE: International Reading Association.
- Rasinski, T., Blachowicz, C., & Lems, K. (Eds.). (2006). *Fluency instruction: Research-based best practices*. New York: The Guilford Press.
- Rayner, K., Chace, K. H., Slattery, T. J., & Ashby, J. (2006). Eye movements as reflections of comprehension processes in reading. *Scientific Studies of Reading, 10*, 241–255.
- Rayner, K., & Sereno, S. C. (1994). Regressive eye movements and sentence parsing: On the use of regression-contingent analyses. *Memory and Cognition, 22*, 281–285.
- Samuels, S. J. (2002). Reading fluency: Its development and assessment. In A. E. Frastrup & S. J. Samuels (Eds.), *What research has to say about reading instruction* (3rd ed.) (pp. 166–183). Newark, DE: International Reading Association.
- Samuels, S. J., & Farstrup, A. E. (2006a). Reading fluency instruction: Will it be a passing fad or permanent fixture? In S. J. Samuels & E. Farstrup (Eds.), *What research has to say about fluency instruction* (pp. 1–3). Newark, DE: International Reading Association.
- Samuels, S. J., & Farstrup, A. E. (Eds.). (2006b). *What research has to say about fluency instruction*. Newark, DE: International Reading Association.
- Schreiber, P. A. (1987). Prosody and structure in children's syntactic processing. In R. Horowitz & S. J. Samuels (Eds.), *Comprehending oral and written language* (pp. 243–270). San Diego: Academic Press.
- Schreiber, P. A. (1991). Understanding prosody's role in reading acquisition. *Theory into Practice, 30*, 3, 158–164.
- Schmitt, N. (Ed.). (2004). *Formulaic sequences*. Amsterdam: John Benjamins.
- Schwanenflugel, P. J., Hamilton, A. M., Kuhn, M. R., Wisenbaker, J. M., & Stahl, S. A. (2004). Becoming a fluent reader: Reading skill and prosodic features in the oral reading of young readers. *Journal of Educational Psychology, 96*, 119–129.
- Schwanenflugel, P. J., Meisinger, E. B., Wisenbaker, J. M., Kuhl, P. K., Strauss, G. P., & Morris, R. D. (2006). Becoming a fluent and automatic reader in the early elementary school years. *Reading Research Quarterly, 41*, 496–522.
- Schwanenflugel, P. J., & Ruston, H. P. (2008). Becoming a fluent reader: From theory to practice. In M. R. Kuhn & P. J. Schwanenflugel (Eds.), *Fluency in the classroom* (pp. 1–16). New York: The Guilford Press.
- Taguchi, E., Takayasu-Maass, M., & Gorsuch, G. J. (2004). Developing reading fluency in EFL: How assisted repeated reading and extensive reading affect fluency development. *Reading in a Foreign Language, 16*, 70–96.

- Topping, K. J. (2006). Building reading fluency: Cognitive, behavioral, and socioemotional factors and the role of peer-mediated learning. In S. J. Samuels & E. Farstrup (Eds.), *What research has to say about fluency instruction* (pp. 106–129). Newark, DE: International Reading Association.
- Yamashita, J. (2007). Investigating asymmetry in EFL learners' mental lexicon: Connections between lexical and conceptual representations in L1 and L2. *JACET Journal*, 45, 63–79.
- Wolf, M., & Katzir-Cohen, T. (2001). Reading fluency and its intervention. *Scientific Studies of Reading*, 5, 211–239.

## Appendix

### *Chunks in the texts*

Basic principles of segmenting texts into chunks (These principles were modified depending on the length of a possible chunk and the connectivity of meanings):

1. Punctuation marks (commas and periods) are marked as chunk boundaries.
2. Syntactic units such as a subject, a verb phrase, an adverbial phrase and a short clause are marked as chunks.
3. Connective devices consisting of multiple words (e.g., “that is” and “for example”) are marked as chunks.

#### Text 1. Leonardo da Vinci's Parachute

Recently, / in the sky above South Africa, / there appeared a strange, white pyramid. / Under this object, / floating 2,000 meters in the air, / hanged the figure of a man. / An Englishman / named Adrian Nicholas / had just become the first person / to use a parachute / designed by / the great Renaissance artist and inventor / Leonardo da Vinci. /

Da Vinci sketched his design / in a notebook / in 1483. / Mr. Nicholas and his partner in the project, / Karen Olsen, / constructed the parachute / using only materials / that would have been available / in Leonardo's time. /

The ride itself / was much smoother and slower / than expected. / Nicholas stated after his achievement, / “All the experts agreed / it wouldn't work. / That is, / it would turn over or / fall apart or / spin around and / make you sick. / But Leonardo was right / all along.” /

#### Text 2. Green Travel

Over the last decade, / well-traveled tourists / have been searching for / more and more new places to visit. / But unfortunately, / promoting tourism in such places / can mean significant damage / to the local environment. /

To help reduce the possible damage, / many tour operators / follow guidelines / suggested by National Environmental Society. / These guidelines promote / a “stay on the path” policy. / This means that / tourists are not allowed / to walk away from the paths / and may not remove any plant or animal life. / Also, / marine life must not be troubled, / and a certain minimum distance / should be kept from wild animals. /

The question is / how many tourists can visit / before a site is permanently damaged. / The key is / to find the right balance / between the money / needed to develop and maintain new tourist areas / and the potential damage to their ecosystems. /

### Text 3. Australian Rules Football

The origin of Australian Rules Football / is unclear. / Some people say / it might have developed / from an ancient game / in which a ball made of kangaroo skin / was kicked around. / Another theory is that / it's related to Ireland's Football. / It is more likely, however, / to be a purely Australian invention of the mid-nineteenth century, / borrowing some rules from rugby. The game was / more or less / in its modern form / by the founding of the Victorian Football Association / in 1877. /

Efforts to export the game to New Zealand and England / failed, / but today / Australian Rules Football / is the country's most popular winter sport. / To an outsider / the game may appear violent and disorderly, / but millions of Australians / simply don't care. / Australian Rules Football will continue / to be a special part / of the country's national identity. /

### Text 4. Moringa — the Tree of Life

Moringa, / a tree that grows / in many nations around the world, / is only now being appreciated / for its wide variety of uses. / It is especially important / in reducing developing countries' dependence / on the industrialized world. / For example, / the cost of cooking oil in African nations / can be double / what we pay at local supermarkets. / If these countries could inexpensively manufacture / their own cooking oil from moringa seeds, / precious funds could be used / to buy other products. /

All of the moringa's uses / are impressive, but / the function with the greatest and / most far-reaching impact / is water cleaning. / Over 6 million children die / every year / from diseases / related to the use of unclean water. / The moringa tree / would reduce / the cost of water treatment. / The moringa is a tree / waiting to be discovered. /

### About the Authors

Junko Yamashita is a professor of language education sciences at Nagoya University, Japan. Her research interest includes L2 literacy (reading) development, L1 influence on L2 processing and acquisition, and assessment of reading ability. E-mail: [yamashita@cc.nagoya-u.ac.jp](mailto:yamashita@cc.nagoya-u.ac.jp)

Shingo Ichikawa is in the PhD program at Carnegie Mellon University, USA. His research interest includes the impact of writing system on reading and the relationship between L1 and L2 reading of bilingual children. E-mail: [sichikaw@andrew.cmu.edu](mailto:sichikaw@andrew.cmu.edu)