

L2 reading fluency progression using timed reading and repeated oral reading

Torrin R. Shimono
Kindai University
Japan

Abstract

The progression of silent reading rates and reading comprehension were examined among Japanese university students ($N = 55$) over one academic semester. Participants were divided into three quasi-experimental groups. The first group practiced a combination of timed reading and repeated oral reading with attention paid to chunking and prosody. The second group practiced timed reading exclusively. The comparison group engaged in oral communication training. Reading rate data were examined using four scoring methods. The results indicated that the treatment groups made statistically significant rate gains, ranging from 13 to 27 standard words per minute. While comprehension percentages were slightly below 70% for both groups, rate gains were accompanied by comprehension increases during the treatment. Both treatment groups outperformed the comparison group in terms of rate and comprehension. While no statistically significant differences were found between the treatment groups, both types of practice were efficacious in promoting second language reading fluency.

Keywords: reading fluency, timed reading, oral reading, repeated reading, reading rate, reading comprehension, chunking, prosody

Being able to read fluently—that is, reading with a high degree of speed and comprehension—is an important skill for second language (L2) learners in both academic and real-world contexts. Reading quickly with sufficient understanding allows L2 learners to process more information in less time and potentially with less effort. This skill can result in higher English proficiency test scores, greater overall L2 academic achievement, and more enjoyment of reading tasks. More practically, in a fast-paced, globalized world connected by the Internet where the majority of webpages are in English (W3Techs, 2017), reading faster enables learners to access a wealth of information. Despite the need for reading fluency, it is a neglected and underdeveloped skill for many L2 learners of English (Grabe, 2009), and Japanese learners are no exception.

In many secondary education contexts in Japan, L2 reading is done for the purpose of learning new vocabulary and grammatical structures, not to build fluency with existing L2 knowledge (Browne, 1998; Gorsuch, 1998). Unfortunately, much intensive L2 reading hardly resembles natural reading in the first language (L1). Instead, it is often a laborious endeavor with large

amounts of reading regressions, translation, and dictionary use, which leads to performance speeds that are three to five times slower than natural L1 reading (Carver, 1990; Chang, 2010). Takase (2003) noted that some Japanese learners read English starting from the end of the sentence in order to make the word order more akin to Japanese syntax. For many L2 learners, this type of reading is tantamount to reading pain—a process where readers exert a large amount of effort in return for limited comprehension and motivation (Extensive Reading Foundation, 2011). Given that L2 reading experts have stressed the pedagogical importance of activities that foster reading fluency, such as reading rate build-up practice as well as text rereading and recycling (Grabe, 2010), the focus of the current study is on the development of silent reading fluency over one academic semester via timed reading and repeated oral reading among Japanese university learners.

Literature Review

Defining and Operationalizing Reading Fluency

Reading fluency is rooted in the automatic recognition of words (LaBerge & Samuels, 1974). It requires well-specified knowledge of a word's phonological, orthographic, and semantic components to be accessed and retrieved accurately from memory while consuming few cognitive resources. This mental process has been described as *ballistic*, where this term refers to the rapid and unstoppable nature of the process (Perfetti & Hart, 2002; Segalowitz, 2003). With a sufficiently automatized L2 mental lexicon that is the product of prolonged practice and task engagement, readers can begin to comprehend a text's meaning accurately at a suitable rate, and the process is perceived to be effortless (Yamashita & Ichikawa, 2010). Thus, Grabe (2009) encapsulated the definition of reading fluency when he said that it is “the ability to read rapidly with ease and accuracy, and to read with appropriate expression and phrasing. It involves a long incremental process, and text comprehension is an expected outcome of fluent reading” (p. 291).

For measurement purposes, reading fluency has often been operationalized by performance time (Breznitz, 2006). In addition, another key variable is sufficient comprehension, and a common method for its assessment is using multiple-choice questions. Nation (2005) stated that for fluent reading, a score of 70% is a benchmark for sufficient comprehension. Thus, previous research has shown that reading rate and comprehension are common and reliable metrics to evaluate silent reading fluency.

Developing Reading Fluency

Reading fluency has been fostered in several ways in reading research. One effective method has been extensive reading, where reading rate and comprehension are increased as a product of substantial time on task and a large amount of words processed (Beglar, Hunt, & Kite, 2012). Reading for pleasure not only allows learners to read faster, but the practice also facilitates a sustainable reading habit. A second effective way to enhance reading rate is repeated reading. Participants read the same passage multiple times during one session to develop automaticity of word recognition. Sometimes, recorded audio support of the reading passages is used during the procedure (Samuels, 1979; Taguchi, Gorsuch, Lems, & Rosszell, 2016). However, the repetitive

nature of the task sometimes has adverse effects on motivation (Taguchi, Gorsuch, Takayasu-Maass, & Snipp, 2012), and rate gains sometimes do not transfer to other texts unless there is a large overlap in vocabulary among the different texts (Rashotte & Torgesen, 1985). A final way to improve fluency is timed reading or speed reading. Learners read under time pressure to increase their reading rate and outperform previous personal benchmarks. This approach has also been shown to be effective at improving performance speeds in many contexts (Chang, 2010; Chung & Nation, 2006; Macalister, 2008; 2010). More recently, oral reading has received attention in the literature as many L1 researchers have argued for prosody's place in assessing reading fluency (National Reading Panel, 2000; Schrauben, 2010). Grabe (2010) supported this notion for L2 contexts by saying, "Oral rereading practice is a good idea and should be used in class" (p. 78). While reading aloud practice is often unpopular among English as a second language and English as a foreign language (ESL and EFL) teachers (Gibson, 2008), it aids in giving Japanese learners a feel for the prosodic and rhythmic features of the target language, especially at the beginning and intermediate levels (Takeuchi, 2003). It also can inspire confidence in the spoken production of English, as it helps attune learners to the stress timing of the language. Oral reading aids in the development of chunking, connected speech, and understanding the relative importance of content words that are important in building fluency (Schreiber, 1987; Yamashita & Ichikawa, 2010). Hence, training in prosody via oral reading is important for the development of reading fluency because it not only is associated with higher comprehension levels but also greater overall reading proficiency (Rasinski, Rikli, & Johnston, 2009).

L2 Reading Fluency Studies

A speed reading course conducted in Korea with 49 first-year university students was shown to be effective in increasing the learners' reading rates (Chung & Nation, 2006). Learners read 23 texts from *Speed Reading* (Quinn & Nation, 1974) over a period of nine weeks. Each text contained 550 words with approximately 12,650 words read in total. The amount of reading practice varied from two to four texts a week. Positive results were found as nearly all participants made rate increases according to three scoring methods—the average rate of the last three passages minus the average of the first three, the highest reading rate passage minus the lowest passages, and the 20th passage minus the first passage. The mean rate of the participants improved by 73 wpm (141 to 214 wpm), 132 wpm (116 to 248 wpm), and 97 wpm (121 to 219 wpm) by these measures, respectively. It should be noted that positive gains are assumed under these scoring methods. However, losses in fluency are possible, especially considering the highest minus the lowest passage, where losses might be seen if the lowest occurs after the highest. The researchers also found that while a majority of the students made gradual increases in reading rate, most of the increases occurred within the first 10 texts, and similar results were found in a replication study (Chung, 2010). The major flaw of the 2006 study was comprehension scores were not reported, so the degree to which the participants understood the readings remains unknown.

Positive effects of a speed reading course among 116 first-year university students majoring in English have also been found in the Vietnamese context (Tran, 2012). Using the same scoring methods as Chung and Nation (2006), the researcher introduced an additional scoring method called the "three extremes scoring method," which takes the average rate of the fastest three

readings minus the average rate of the slowest three readings. The two experimental groups, group A and group B, read the same 20 timed reading passages ($\approx 11,000$ words) from *Asian and Pacific Speed Readings for ESL Learners* (Quinn, Nation, & Millet, 2007) over a period of three months. Groups C and D served as comparison groups in which the participants in group C were English majors and participants in group D took a course from the language center and were not English majors. For group A, mean increases of reading rate of the treatment passages according to the four scoring methods—the average rate of the last three passages minus the first three, 20th minus first, the fastest minus the slowest, and the average of the three fastest minus the three slowest—were 57.00, 61.03, 97.67, and 80.38 wpm, respectively. For group B, the increases were 50.90, 51.03, 87.83, and 73.16 wpm, respectively. Comprehension scores were over 70% on all measures. While the mean reading rates for Groups C and D increased by 10.46 and 19.65 wpm, respectively, on the pre- and posttest, the treatment groups' increases were more substantial. Contrary to Chung and Nation's (2006) findings, Tran found that most participants increased their speed and comprehension in the latter half of the course.

The benefits of timed reading on fluency have also been shown among 84 high-beginner to lower-intermediate level (TOEIC 325–400) Taiwanese university learners (Chang, 2010). To address some of the methodological shortcomings of Chung and Nation (2006), the researcher included a comparison group ($n = 38$) to contrast with the experimental group ($n = 46$). She also gathered data concerning the participants' general perceptions of the treatment's effectiveness using short interviews. The timed reading texts were chosen from *Reading for Speed and Fluency 2* (Nation & Malarcher, 2007b), with each text containing approximately 300 words. Timed reading practice was done once a week for 13 weeks. In each session, the students in the experimental group spent 15 minutes reading three passages for a total of 39 texts ($\approx 11,700$ words). The comparison group spent time reviewing course content that mainly related to the TOEIC. Using repeated-measures ANOVA with between-subjects contrasts, results showed that the timed reading group increased their reading speed on average by 29 wpm, from 118 to 147 wpm. The comparison group increased by only 7 wpm, from 124 to 131 wpm. The difference between the two within-subjects time periods for the experimental group was statistically significant ($p < .0005$), with a small effect size ($d = .80$) according to the effect size guidelines outlined in Plonsky and Oswald (2014). The within-subjects measure for the comparison group was not statistically significant. However, the between-subjects test indicated no significant difference between the groups. Despite the improvement in reading rates, comprehension levels improved only marginally for both groups and fell below the 70% threshold. From Time 1 to Time 2, the comprehension percentages were 62.54% and 67.28%, respectively, for the experimental group and 60.30% and 64.77%, respectively, for the comparison group. A majority of the interviewees in the treatment group felt that the timed readings helped increase their reading speed, concentration while reading, and general reading confidence in English.

Further reading fluency research in Taiwan revealed that a treatment of timed reading was more effective in producing faster reading rates when compared to a repeated oral reading treatment (Chang, 2012). This experiment was conducted over 13 weeks with 35 low- to intermediate-proficiency adult students (self-reported TOEIC 450 to 550). *Reading for Speed and Fluency 2* and 3 (Nation & Malarcher, 2007b, 2007c) were used for the timed readings. The timed reading group gained about 50 wpm after reading 52 passages silently (39 in-class, 13 out of class—approximately 16,800 words). The repeated reading group, who had no time pressure placed on

them, gained about 23 wpm after reading 26 passages (13 in-class, 13 out-of-class), up to five times in various manners such as reading with audio support, silent reading, individual oral reading, paired oral reading, and volunteer reading. This amounted to an estimated 31,200 words processed. Both groups decreased slightly in reading rate on the delayed posttest as the timed reading group had a loss of 5 wpm and the repeated reading group fell back 4 wpm. The two groups had similar comprehension scores that were below the 70% threshold, much like the findings in her 2010 study. On the pretest, the comprehension percentages for both groups were 53.33%. The posttest percentage was 66.66% for the timed reading group and 60.00% for the repeated reading group. On the delayed posttest, they were 63.33% and 53.33% for the timed reading and repeated reading groups, respectively. The results of the study were understandable because the researcher admitted that the repeated reading group had no incentive to read faster, as no time pressure was placed on them. This study also had control issues. A little more than 25% of the passages for the timed reading group were assigned as outside reading. For the repeated reading group, the number of times participants read varied in each session. Moreover, the way the researcher had participants record reading times of the passages could have been improved. An assistant stood at the front of the classroom flipping pages of a spiral notebook that showed five-second intervals of reading times. Using this method might have caused measurement inaccuracies due to the approximations as well as the time it takes for the participants to look up, identify, and record their times. The study also lacked a comparison group to which the effects of the two treatment groups could be compared.

The differential effectiveness of timed reading and timed repeated reading in improving reading rates and comprehension was compared among 26 Taiwanese university students over a 13-week period (Chang & Millet, 2013). Half of the participants practiced timed reading with 26 passages from *Reading for Speed and Fluency 1* (Nation & Malarcher, 2007a), with approximately 300 words per passage. The repeated reading group read each passage five times ($\approx 39,000$ words read) and answered comprehension questions after the first and the fifth reading. The timed reading participants read the same number of passages but with no reading repetition. Hence, they read each passage only once and answered the comprehension questions after the first reading ($\approx 7,800$ words read). All students were tested based on two practiced texts and one unpracticed text before and after the intervention. The results showed that the repeated reading participants increased their reading rate by 47 wpm (103 to 150 wpm) and 45 wpm (102 to 147 wpm) for the practiced and unpracticed texts, respectively. The timed reading group, or non-repetition participants, increased by 13 (107 to 120 wpm) and 7 wpm (102 to 109 wpm). Repeated reading students' comprehension improved by 19% (51 to 70%) and 17% (49 to 66%) for the practiced and unpracticed texts, while the non-repeated reading students improved by 5% (49 to 54%) and 3% (44 to 47%), respectively. While the results look promising for repeated reading, the small sample size calls into question the generalizability of the results. Additionally, time on task, as well as the number of words processed, appeared to be unequal between the two groups. The researchers also reported that the results suffered from a selection bias in that participants in the repeated reading group appeared more highly motivated than those in the timed reading group.

While few researchers have examined reading fluency treatments in Japan, a six-month speed reading course in a Japanese high school has been shown to improve reading rate (Underwood, Myskow, & Hattori, 2012). The participants were 105 10th-grade students attending a high

school that was ranked slightly below average in terms of its comparative standardized rank score (*hensachi* in Japanese; see Newfields, 2006 for a discussion). Fifty-one students were randomly assigned into an experimental group and 54 students were assigned to a comparison group. Using *Reading for Speed and Fluency 1* (Nation & Malarcher, 2007a), the experimental group received timed reading training twice a week (once in each bi-weekly English class), and the comparison group received supplementary activities focused on high-frequency vocabulary development. The results indicated that both the experimental and comparison groups made significant within-group improvements in general reading comprehension and knowledge of high-frequency vocabulary. However, no significant between-group differences were found except on reading rate. While the experimental group made gains in reading rate by decreasing the reading time by an average of 47 seconds (first three readings: $M = 173$ seconds; last three readings: $M = 125$ seconds, or 104 wpm to 144 wpm), the average rate of comprehension was low, averaging about three out of five multiple-choice questions correct. The results also indicated that other reading processes, such as lexical and grammatical knowledge, are necessary to improve comprehension in addition to vocabulary and word recognition.

In sum, many of the studies reviewed above have demonstrated the effectiveness of longitudinal fluency treatments utilizing timed reading, repeated reading, and oral reading in various contexts. Particularly, substantial reading rate gains have been observed. However, the reviewed studies have methodological problems that are in need of improvement. First, reading rate has rarely been measured in terms of standard words per minute as espoused by Beglar and Hunt (2014). According to Carver (1990), a standard word is made up of six letter spaces of text, which can include letters, punctuation, and spacing. Because words have varying lengths, there is a need to standardize the measurement to make reading rate results across studies more comparable. In addition, Flesch-Kincaid readability statistics are not reported, again making it difficult to compare reading rate across varying difficulties of text. Moreover, task acclimation procedures for the timed reading activities were not performed prior to the treatment. This step is necessary when administering timed readings because often the first and last passages are compared and the first passage reading speed could be artificially low due to unfamiliarity to the task when the participants begin the treatment, thereby skewing the results. Finally, almost all studies mentioned fail to report time on task or explicitly state the number of words read during the treatment; thus, important details of the treatments remain unreported.

This study is designed to address some of the past methodological inadequacies of prior studies, as well as three gaps in the reading fluency literature. First, the effects of two reading fluency treatments—(a) timed reading and (b) timed reading plus oral rereading and chunking practice—have yet to be explored on reading rate in the Japanese university context. Second, the effects of these treatments on reading comprehension also have not been examined. Third, the effects of these reading treatments have not been compared. Thus, the primary purpose of this study is to provide a detailed account of the progression of silent reading rates over one academic semester among lower proficiency Japanese university students who receive treatments of timed reading, as well as the combination of timed reading and oral reading. Four scoring methods outlined in Chung and Nation (2006) and Tran (2012), along with statistical significance testing, were employed to test reading rate. The second purpose is to illustrate comprehension progress of participants over the treatment period, as well as statistically test the comprehension rates of the

passages that were identified under the four scoring methods. The final purpose is to compare the efficacy of the two fluency treatments along with a comparison group.

This study addresses three research questions (RQs):

- RQ1. To what extent do the reading fluency treatment groups increase their reading rates over one academic semester?
- RQ2. To what extent do the reading fluency treatment groups increase reading comprehension over one academic semester?
- RQ3. How do the reading fluency treatment groups, along with a comparison group, compare with each other in terms of reading rate and reading comprehension by the end of the treatment period?

Methodology

Participants

Initially, 60 students agreed to participate in the study; however, five participants were removed from the data analysis due to extended absences. Thus, 55 first- and second-year Japanese students (30 males, 25 females; 33 first-year, 22 second-year) from a lower tier Japanese university (*hensachi* = 38) in western Japan participated in this study. Thirty-eight participants in the treatment groups were members of the Faculty of Foreign Languages. Seventeen participants from the Faculty of Law formed the comparison group. Foreign language majors had five required 90-minute English classes per week, while law students had two 90-minute English classes per week. Their ages ranged from 18 to 20 years old.

The foreign language majors were divided into two treatment groups: One received a combination of timed reading and repeated oral reading training with chunking practice ($n = 20$, henceforth the TROR group) and the other practiced timed reading only ($n = 18$, henceforth the TR group) (see Table 1). Both groups consisted of two intact classes with one first-year and one second-year class each. The researcher, who is North American, taught both classes from the TR group, while two other North American instructors taught the TROR group classes. The comparison group ($n = 17$) was also made up of two intact first-year classes that focused on speaking, listening, and communication. Both classes were taught by the researcher.

The Faculty of Foreign Languages streamed classes into three proficiency levels—lower, middle, and upper levels. The participants from the TROR group were members of classes designated as lower level, and the participants of the TR group were members of the middle proficiency band. For the comparison group, one class was designated as lower and the other was middle proficiency. TOEIC Bridge scores of the students from the lower level ranged from 90 to 110 (230 to 280 TOEIC equivalent), while middle level student scores ranged from 120 to 140 (310 to 395 TOEIC equivalent) (Educational Testing Service, 2006).

Table 1. *Treatment, English proficiency, and written receptive vocabulary size of TROR (n = 20), TR (n = 18), and comparison groups (n = 17)*

Group	Treatment	Class and Proficiency	Vocabulary Size
TROR	<ul style="list-style-type: none"> • Timed reading (2x per week) • 4 oral readings (2 choral and 2 individual per week) 	First-year: Lower Second-year: Lower	1,815
TR	<ul style="list-style-type: none"> • Timed reading (3x per week) 	First-year: Middle Second-year: Middle	1,988
Comparison	<ul style="list-style-type: none"> • Speaking and listening practice 	2 first-year classes: Lower and Middle	1,800

Vocabulary Size Test

Nation and Beglar's (2007) Vocabulary Size Test was utilized to gauge participants' written receptive vocabulary knowledge in order to select appropriate reading materials for the study. To avoid test fatigue, only the first through fourth 1,000 word frequency levels of the test were administered to the participants at the beginning of the semester. The test has 10 items per word frequency level (10 items x 4 word frequency levels = 40 items). The mean estimated vocabulary size score was 1,815 for the TROR group, 1,989 for the TR group, and 1,800 for the comparison group. A one-way ANOVA revealed no significant differences between the groups in terms of their vocabulary size scores $F(2, 52) = 1.82, p = .17$.

Treatment Materials

Based on the results of the Vocabulary Size Test, it was decided that *Reading Power: Reading for Pleasure, Comprehension Skills, Thinking Skills, Reading Faster*, third edition, (Mikulecky & Jeffries, 2005) would be used for the timed reading materials because they used vocabulary predominantly from the first 1,000 most frequent words of English, which was suitable for the participants' reading proficiency level. This set of timed readings is about the lives of an American family and their experiences. These readings included 20 passages with a set of comprehension questions for every passage. For the TR group, the first 10 timed readings from Quinn, Nation, and Millet's (2007) speed reading course, *Asian and Pacific Speed Readings for ESL Learners* served as supplementary readings for their treatment. These passages were also written at the 1,000 word frequency level and were focused on topics related to Asia, ranging from death ceremonies in Bali to Buddhism.

The readability statistics of these passages, which include Flesch Reading Ease, Flesch-Kincaid Grade Level, percentage of passive sentences, number of words, and number of comprehension questions are provided in Table 2. The Flesch Reading Ease is a formula that quantifies the difficulty of the passage by taking into account the total number of sentences, total number of words, and total number of syllables. The higher the number, the easier a passage is to read with the highest score being 100.00. The timed reading passages in this study ranged from 71.8 to 86.7. The Flesch-Kincaid Grade Level is another readability formula that estimates the difficulty of the passages which also takes into consideration the total number of sentences, total number of words, and total number of syllables. The resulting score is a number that corresponds with a U.S. grade level, or the number of years of formal education required to understand a text. Thus, the

timed reading passages used in this study ranged from 3.1 to 5.7, indicating that the difficulty of the passages ranged from approximately a third to sixth grade reading level.

Timed Reading Pre- and Posttest

Two timed reading warm-up passages from *Reading Power* were used as pre- and posttest measures to assess changes in reading rate and comprehension. The readability statistics and passage characteristics are shown in Table 2. The two passages were counterbalanced and randomly distributed to participants at the beginning of the semester. Thus, if a participant received passage A for the pretest, passage B was given as the posttest and vice versa.

Table 2. *Readability statistics, number of words, and number of comprehension questions for the pretest, posttest, and treatment passages*

	Flesch Reading Ease	Flesch-Kincaid Grade Level	Percentage of passive sentences	Approximate number of words	Number of comprehension questions
Pre- and Posttest	71.8	5.7	0	200	8
<i>Reading Power</i> : Passages 1–15	86.7	3.1	0	200	8
<i>Reading Power</i> : Passages 16–20	84.9	3.6	0	400	8
<i>Asian and Pacific Speed Readings</i> : Passages 1–10	78.9	5.7	10	550	10

Procedure

This study was conducted over 12 weeks during the spring semester at a Japanese university. In the first session of the first week, instructors explained the nature of the study, obtained consent forms, and administered the Vocabulary Size Test. For the comparison group, the reading pretest was given in the first session. In the second session of the first week, participants in the treatment groups were informed about reading fluency strategies in English; however, some Japanese was used for clarification. Participants were told about the benefits of reading faster. By doing so, it would make reading more enjoyable and comprehensible. They were encouraged not to translate every word, skip unknown words, and chunk thought groups together instead of reading one word at a time. Furthermore, they were told that good readers vary their reading speed as some sentences are read faster and some slower, and occasionally, it is necessary to re-read certain parts in order to ensure adequate comprehension (Mikulecky & Jeffries, 2005).

Following the strategy training, the participants were given two practice timed reading passages in order to acclimate them to the task. Data from these practice passages were not used in this study. Instructors also informed the participants that they should aim for a minimum of a 70% comprehension rate. Thus, if students' comprehension fell below 70%, they were told to read more carefully. In contrast, if a perfect score was achieved, they were told to push themselves to read faster. In the third session of the first week, the timed reading pretest was given to all the treatment group participants. Over the next 10 weeks, the timed reading passages were

administered to the treatment groups. In week 12 of the semester, the timed reading posttest was given to all participants.

Timed Reading Treatment

Both the TROR and TR groups received timed reading training. The TR group received three timed reading passages a week on separate days. Because the TROR group received additional repeated oral reading and chunking practice, only two reading passages were given on different days in order to balance time on task. Thus, the TR group read 30 timed readings (10,620 words or 9,819 standard words) and the TROR group read 20 timed readings (5,120 words read three times for a total of 15,360 words or 13,841 standard words). If participants were absent, they were instructed to read previously unread passages in the next class session.

For the timed readings, every participant was given an individual stopwatch for measurement accuracy. The instructors told the students to try to read as fast as they could but not at the expense of losing comprehension. Moreover, they were continuously encouraged to increase their reading rate from the previous passage. For the comprehension questions, their goal was to answer at least six out of the eight questions correctly for the *Reading Power* passages and seven out of 10 for the *Asian and Pacific Speed Reading* passages in order to achieve the 70% criterion.

After the instructor distributed the passages, the participants were instructed to press the start button on their stopwatch and begin reading the passage. After finishing, they pushed the stop button and recorded their reading times at the bottom of the page, exactly as indicated on the stopwatch. The participants then answered the comprehension questions on the other side of the page without looking back at the passage. Instructors carefully monitored the students so that they would not turn their page over to reread sections. In addition, all comprehension questions were printed on the back, upside-down to the text of the reading passage page to prevent participants from obtaining answers that could be potentially seen through the page. After the comprehension questions were finished, they were corrected in class. Reading rates and comprehension scores were then recorded on the participants' graphs, which allowed them to keep track of their progress. Each individual session for the TR group took approximately 10 minutes. Because there were 30 sessions for the TR group, the approximate time on task was 300 minutes over 10 weeks.

Repeated Oral Reading with Chunking and Prosody Treatment

For the TROR group, after the participants completed the timed reading passage, the instructor handed out the same timed reading passage they had just read, except it was marked with forward slashes that separated 3–5 word phrases and thought groups (see the Appendix for a sample reading). They also indicated to the participants where they should try to chunk groups of words. The instructor then read the passage aloud, pausing after each chunk. The participants then repeated the same chunk aloud while reading it, trying to mimic the instructor's prosodic qualities, as well as the pace and rhythm of the reading. After one choral reading with the instructor, the participants practiced reading the passage aloud once again to a partner, trying to read the passage as fluently as possible, with attention being paid to chunking and prosody. Thus, each passage was read a total of three times—once silently and twice orally. Each treatment took

about 15 minutes. Because there were 20 sessions for the TROR group, the approximate time on task was 300 minutes for the entire treatment.

Comparison Group

The comparison group had two English classes per week geared toward developing speaking and listening skills, as well as communicative competence. While there were eight short reading passages in the class textbook read over the semester ($\approx 1,501$ words or 1,335 standard words), no explicit reading fluency treatment was given to these participants.

Analyses

To analyze the results of the treatment data, reading fluency gains were operationalized as reading rate gains made while maintaining comprehension levels. In order to ascertain the participants' reading rates, first, the number of all characters and spaces were counted in all the reading passages. Then, this sum was divided by 6 in order to calculate the number of standard words per passage. Next, the number of standard words of each passage was divided by the total time each participant took to complete the passage in seconds and multiplied by 60. Thus, the resulting values reflected the participants' reading rates in terms of standard words per minute (swpm). Comprehension scores were calculated into percentages out of 100.

The changes in the reading rates and comprehension percentages of the timed reading passages were analyzed quantitatively through SPSS (version 24). Because multiple *t* test comparisons were carried out, a sequential Holm-Bonferroni adjustment (Holm, 1979) was employed to help control for Type I errors. Effect sizes were evaluated using Cohen's *d*, where 0.60 is considered a small effect size, 1.00 is medium, and 1.40 is a large effect for within-group contrasts and .40, .70, and 1.00 constitutes small, medium, and large effect sizes for between-subject contrasts (Plonsky & Oswald, 2014).

Results

RQ1 seeks to investigate to what extent the reading fluency treatment groups increased their reading rates over one academic semester. To answer the first research question, descriptive statistics of the groups' reading rate progression by week are reported, along with pre- and posttest measures (see Table 3). The figures reported during the treatment period represent the average reading rate and comprehension percentage of the treatment groups for each given week. Hence, for the TROR group, two passages were averaged because they read two passages per week, and three passages were averaged for the TR group because they read three passages per week. Additionally, the treatment period was analyzed by dividing the treatment passages into three phases—the first three weeks (weeks 2 to 4), middle four weeks (weeks 5 to 8), and final three weeks (weeks 9 to 11)—and the average reading rates for each phase were calculated. Finally, the gains in reading rate were tested for statistical significance utilizing four scoring methods and effect sizes are reported.

Table 3. Means, standard deviations, and 95% confidence intervals of reading rate in standard words per minute for the TROR, TR, and comparison group

	TROR Group			TR Group			Comparison Group		
	<i>M</i>	<i>SD</i>	95% <i>CI</i>	<i>M</i>	<i>SD</i>	95% <i>CI</i>	<i>M</i>	<i>SD</i>	95% <i>CI</i>
Pretest (Week 1)	83.76	28.04	[70.71, 96.89]	93.54	22.53	[82.28, 104.72]	101.53	29.23	[86.50, 116.56]
Week 2	83.00	21.35	[73.01, 92.99]	95.80	23.52	[84.10, 107.49]			
Week 3	85.75	21.06	[75.89, 95.61]	106.06	33.38	[89.46, 122.65]			
Week 4	91.35	25.17	[79.57, 103.13]	105.56	29.03	[91.12, 119.99]			
Week 5	83.60	25.61	[71.62, 95.58]	98.22	26.75	[84.92, 111.52]			
Week 6	89.95	30.77	[75.55, 104.35]	109.06	23.59	[97.33, 120.78]			
Week 7	93.35	33.25	[77.79, 108.91]	98.46	25.04	[86.01, 110.92]			
Week 8	94.80	30.27	[80.63, 108.97]	105.59	29.57	[90.89, 120.30]			
Week 9	96.65	24.48	[85.19, 108.11]	107.09	25.69	[94.32, 119.87]			
Week 10	104.25	37.07	[86.90, 121.60]	108.20	24.98	[95.78, 120.63]			
Week 11	98.95	27.78	[85.95, 111.95]	112.76	19.06	[103.28, 122.24]			
Posttest (Week 12)	96.34	35.19	[79.90, 112.90]	108.96	19.19	[99.45, 118.55]	85.47	33.31	[68.34, 102.60]

Figure 1 shows the reading rate progression of the groups in terms of week of the treatment. For both groups, the first half of the treatment shows some progressions and regressions. However, there is an upward trend toward the end of the treatment period for both groups. Looking at the treatment period in the three phases—the beginning, middle, and end—the mean reading rates for the first six (first three weeks), middle eight (middle four weeks), and final six passages (last three weeks) for the TROR group were 86.50, 90.25, and 99.67 swpm, respectively. For the TR group, the mean rates were 102.55, 103.00, and 109.44 swpm for the first nine, middle 12, and final nine passages, respectively, confirming upward trends throughout the treatment period for both groups. The mean reading rate for all 20 passages was 91.95 swpm for the TROR group, while the mean rate for all 30 passages was 105.07 swpm for the TR group. For the comparison group, the initial reading rate was 101.53 swpm, which was the highest initial group reading rate. However, the final rate was 85.47 swpm, showing a statistically significant decrease in reading rate from pre- to posttest measures, $t(16) = 3.05$, $p = .008$, $[4.89, 27.23]$, $d = 0.74$.

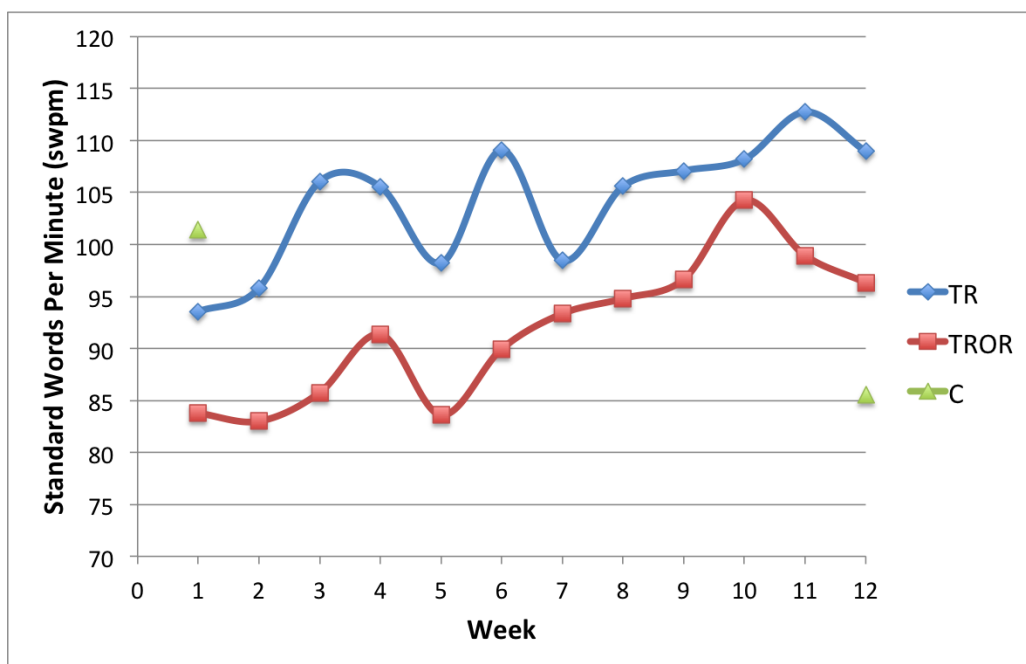


Figure 1. 12-week reading rate progression of the TROR, TR, and comparison groups

To test for statistical significance of the changes in reading rate for the treatment groups, paired-samples t tests were conducted to determine within-subjects changes in reading rate using four scoring methods. The *average scoring method* compares the average reading rate of the first three passages in the treatment and the last three passages (TROR: passages 1–3 vs. passages 18–20; TR: passages 1–3 vs. passages 28–30). The *last minus the first scoring method* compares the pre- and posttest reading rates. The *extreme scoring method* contrasts the passage with the fastest rate and the passage with the slowest rate. The *three extremes scoring method* compares the average of the fastest three passages and the slowest three passages. Assumptions of the t tests were checked and met. In this method, the alpha value, set at $p = .05$, was divided sequentially by the number of tests performed. Because there were four comparisons for each group, significance of the results was reached when the p values obtained were ordered sequentially

from smallest to largest and compared to the adjusted alpha level of $p \leq .013$, $.017$, $.025$, and $.05$, respectively.

Table 4 shows the reading rates, gain scores, and statistical significances of the four scoring methods of the TROR group. It is apparent that the average and last minus first methods have closer mean rates for the comparative measures, while the extreme and three extremes means are closer to each other. Evaluating the TROR group in terms of the average scoring method showed a significant reading rate increase of about 15 swpm, from 84 to 99 swpm, $t(19) = -3.18$, $p = .005$. The standardized effect size index, Cohen's d , was -0.71 , which constitutes a small effect size. For the last minus the first scoring method, the gain was about 13 swpm, $t(19) = -2.79$, $p = .012$. Cohen's d was -0.62 , also indicating a small effect. The fastest passage for the TROR group was passage 17 in which the group average was 107.90 swpm, while the slowest was passage 2 and the group mean was 80.55 swpm. Hence, the extreme scoring method showed the largest rate gain of 27 swpm, $t(19) = -2.52$, $p = .021$, with a small effect size, $d = -0.56$. Finally, the three fastest passages were 12 (100.50 swpm), 17 (107.90 swpm), and 20 (101.40 swpm), and the three slowest were passages 2 (80.55 swpm), 6 (83.40 swpm), and 8 (83.05 swpm). Thus, because the three slowest passages occurred prior to the three fastest, the three extreme method showed a positive gain of about 21 swpm, $t(19) = -2.96$, $p = .008$ with a small effect, $d = -0.66$. When ordered sequentially from lowest to highest, all p values reached statistical significance according to the Holm-Bonferroni procedure.

Table 4. *Reading rate means, standard deviations, 95% confidence intervals, and gains of the four scoring methods for the TROR group (n = 20)*

Scoring method	<i>M</i>	<i>SD</i>	95% <i>CI</i>
Average			
First 3	83.80	21.00	[73.97, 93.63]
Last 3	99.08	26.18	[86.83, 111.33]
Gains	15.28**	21.49	[5.22, 25.34]
Last minus first			
Pretest	83.76	28.04	[70.71, 96.89]
Posttest	96.34	35.19	[99.45, 118.55]
Gains	12.58*	20.18	[3.14, 22.02]
Extreme			
Slowest	80.55	22.61	[69.97, 91.13]
Fastest	107.90	51.04	[84.01, 131.79]
Gains	27.35*	48.56	[4.62, 50.08]
3 Extremes			
Slowest 3	82.60	22.47	[72.08, 93.12]
Fastest 3	103.27	35.41	[86.70, 119.84]
Gains	20.67**	31.26	[6.04, 35.30]

Note. First 3 = first three passages of the treatment. Last 3 = last three passages of the treatment. Gains = increase in reading rate in swpm. Last = posttest. First = pretest. Slowest = slowest passage of the treatment. Fastest = fastest passage of the treatment. Slowest 3 = slowest three passages of the treatment. Fastest 3 = fastest three passages of the treatment.

* = $p \leq .05$. ** = $p \leq .01$

Table 5 shows the reading rates, gain scores, and statistical significances of the four scoring methods for the TR group. Like the TROR group, the final means of the extreme and three extreme scoring methods are higher than the means of the average and last minus the first methods. The average scoring method showed a rate gain of approximately 17 swpm, from 96 to 113 swpm, $t(17) = -2.68$, $p = .016$. The effect size was small at -0.63 . The last minus the first scoring method saw a gain of about 15 swpm, $t(17) = -4.26$, $p = .001$, and a medium effect size ($d = -1.01$). For the extreme scoring method, the passage with the highest mean rate was 15 (120.17 swpm) and the lowest was the pretest (93.54 swpm). Thus, the group gain was about 27 swpm, $t(17) = -3.35$, $p = .004$, with a small effect ($d = -0.80$). For the three extremes method, the three fastest passages were 15, 28, and 29 and the means were 120.17, 112.94, and 116.67 swpm, respectively. The three slowest were the pretest passage, along with passage 2 and 10, which were read at 93.54, 93.78, and 94.56 swpm, respectively. Because the three fastest occurred after the three slowest, there was a positive increase of approximately 21 swpm, $t(17) = -3.92$, $p = .001$ with a small effect size ($d = -0.92$). Using the Holm-Bonferroni procedure, all p values reached statistical significance.

Table 5. Reading rate means, standard deviations, 95% confidence intervals, and gains of the four scoring methods for the TR group ($n = 18$)

Scoring method	<i>M</i>	<i>SD</i>	95% <i>CI</i>
Average			
First 3	95.80	23.52	[84.10, 107.49]
Last 3	112.76	19.06	[103.28, 122.24]
Gains	16.96*	26.81	[3.63, 30.30]
Last minus first			
Pretest	93.54	22.53	[82.28, 104.72]
Posttest	108.96	19.19	[100.89, 116.44]
Gains	15.42**	15.36	[7.78, 23.05]
Extreme			
Slowest	93.54	22.53	[82.28, 104.72]
Fastest	120.17	34.81	[102.85, 137.48]
Gains	26.63**	33.81	[9.85, 43.48]
3 Extremes			
Slowest 3	93.96	20.51	[83.76, 104.16]
Fastest 3	116.59	24.19	[104.56, 128.62]
Gains	22.63**	24.48	[10.46, 34.81]

Note. First 3 = first three passages of the treatment. Last 3 = last three passages of the treatment. Gains = increase in reading rate in swpm. Last = posttest. First = pretest. Slowest = slowest passage of the treatment. Fastest = fastest passage of the treatment. Slowest 3 = slowest three passages of the treatment. Fastest 3 = fastest three passages of the treatment.

* = $p \leq .05$. ** = $p \leq .01$

Table 6. Means, standard deviations, and 95% confidence intervals of comprehension percentages for the TROR, TR, and comparison group

	TROR Group			TR Group			Comparison Group		
	<i>M</i>	<i>SD</i>	95% <i>CI</i>	<i>M</i>	<i>SD</i>	95% <i>CI</i>	<i>M</i>	<i>SD</i>	95% <i>CI</i>
Pretest (Week 1)	56.25	17.91	[47.87, 64.63]	56.94	20.21	[46.89, 67.00]	53.68	20.14	[43.32, 64.03]
Week 2	62.81	16.16	[55.25, 70.37]	61.11	18.96	[51.68, 70.54]			
Week 3	67.50	20.64	[57.84, 77.16]	72.22	18.69	[62.93, 81.52]			
Week 4	65.94	17.73	[57.64, 74.24]	66.90	13.89	[59.99, 73.81]			
Week 5	59.06	18.42	[50.54, 67.68]	69.68	15.77	[61.83, 77.52]			
Week 6	61.56	15.61	[54.26, 68.87]	69.21	12.39	[63.05, 75.37]			
Week 7	63.44	14.80	[56.51, 70.36]	67.87	18.01	[58.87, 76.78]			
Week 8	77.81	13.06	[71.70, 83.92]	59.86	13.53	[53.13, 66.59]			
Week 9	72.19	16.16	[64.63, 79.75]	58.70	17.00	[50.25, 67.16]			
Week 10	70.94	14.80	[64.01, 77.86]	66.30	13.52	[59.57, 73.02]			
Week 11	68.13	20.47	[58.55, 77.70]	72.41	19.36	[62.78, 82.04]			
Posttest (Week 12)	68.13	13.13	[61.69, 74.27]	70.83	15.46	[63.15, 78.52]	63.97	19.20	[54.10, 73.84]

RQ2 asked to what extent the reading fluency treatment groups increase reading comprehension over one academic semester. Table 6 shows the descriptive statistics of the groups' comprehension scores for the pre- and posttest and weekly averages. Comprehension percentage scores were also analyzed in terms of the beginning, middle, and final passages. Because the treatment period was 10 weeks, the first section was defined as the first three weeks (TROR: passages 1–6; TR: 1–9), the middle consisted of week four to seven (TROR: passages 7–14; TR: 10–21), and the final section was the last three weeks (TROR: passages 15–20; TR: 22–30). Descriptive statistics for each section are reported and analyzed to show the progression of comprehension rates.

Figure 2 shows the progression of comprehension rates over one semester by week. In most instances, both groups fell below the 70% threshold, although the TROR group saw comprehension gains above the threshold toward the end of the treatment. Dividing the passages into the first six, middle eight, and final six passages of the treatment for the TROR group, the comprehension averages were 65.42%, 65.47%, and 70.42%, respectively. For the TR group, the passages were divided by the first nine, middle twelve, and final nine passages, and the averages were 66.74%, 66.64%, and 65.80%, respectively. The average comprehension rate for all 20 passages was 66.94% for the TROR group, while the average comprehension for all 30 passages was 67.07% for the TR group. For the comparison group, the comprehension percentage for the first passage was 53.68% and 63.97% for the final passage and was statistically significant, $t(16) = -3.00, p = .008, [3.03, 17.56], d = -0.73$.

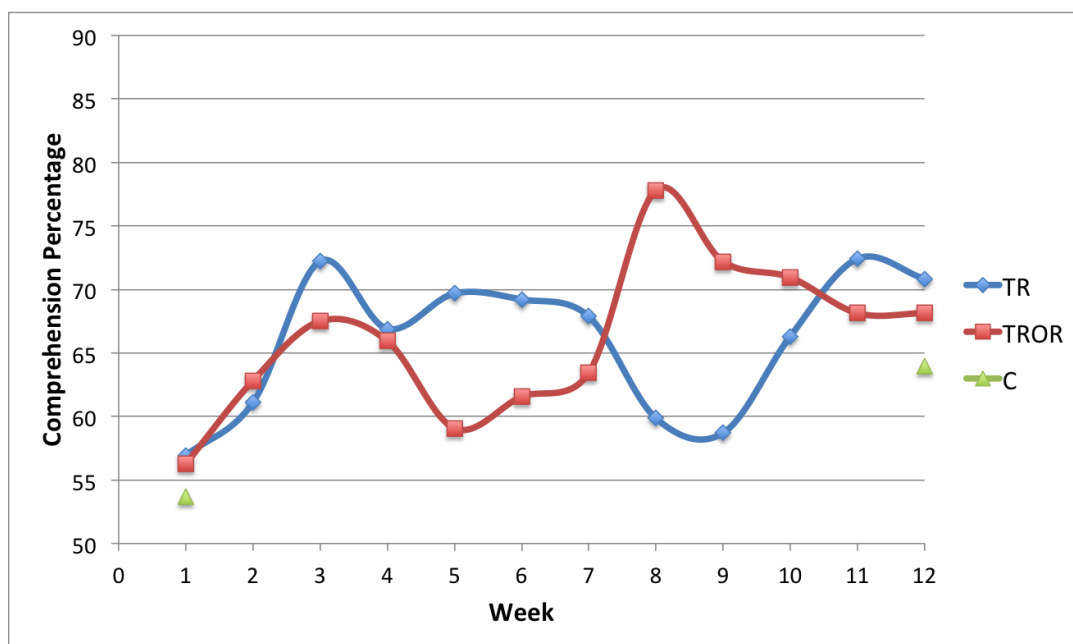


Figure 2. 12-week reading comprehension progression of the TROR, TR, and comparison groups

The comprehension percentages of the passages from the four scoring methods of reading rate were also tested for statistical significance using paired-samples t tests and the Holm-Bonferroni procedure. Table 7 shows the comprehension percentages, gain scores, and statistical significances of the TROR group. The average method passages and last minus first method passages show larger comprehension increases than the extreme method and three extreme

method passages. Comparing the average of the first and last three passages of the treatment, a significant 8% increase in comprehension was found, $t(19) = -2.66, p = .016$. Cohen's d was -0.59 , indicating a small effect size. Contrasting the pre- and posttest comprehension scores, a significant gain of about 12% was observed, $t(19) = -2.97, p = .008$. Cohen's d was -0.66 , which also indicates a small effect. According to the extreme scoring method passages, there was a 9% drop in comprehension between the fastest and slowest passage; however, this difference was not statistically significant, $t(19) = 2.05, p = .054, d = 0.56$. Finally, for the three extreme method passages, a 3% gain in comprehension was observed, but it missed statistical significance, $t(19) = -1.05, p = .309, d = -0.23$.

Table 7. *Comprehension percentage means, standard deviations, 95% confidence intervals, and gains of the four scoring methods for the TROR group (n = 20)*

Scoring method	<i>M</i>	<i>SD</i>	<i>95% CI</i>
Average			
First 3	63.75	17.32	[55.65, 71.85]
Last 3	71.88	17.77	[63.56, 80.19]
Gains	8.13*	13.68	[1.72, 14.53]
Last minus first			
Pretest	56.25	17.91	[47.87, 64.63]
Posttest	68.13	13.13	[61.69, 74.27]
Gains	11.88**	17.90	[3.50, 20.25]
Extreme			
Slowest	71.25	16.77	[63.40, 79.10]
Fastest	62.50	19.45	[53.40, 71.60]
Gains	-8.75	19.07	[-17.67, 0.17]
3 Extremes			
Slowest 3	60.83	16.96	[52.90, 68.77]
Fastest 3	63.54	14.17	[56.91, 70.17]
Gains	2.71	15.37	[-4.49, 9.90]

Note. First 3 = first three passages of the treatment. Last 3 = last three passages of the treatment. Gains = increase in comprehension percentage. Last = posttest. First = pretest. Slowest = slowest passage of the treatment. Fastest = fastest passage of the treatment. Slowest 3 = slowest three passages of the treatment. Fastest 3 = fastest three passages of the treatment.

* = $p \leq .05$. ** = $p \leq .01$

Table 8 illustrates the comprehension percentages, gain scores, and statistical significances for the TR group. The average method passages and last minus first method passages show comparative means that are more similar to each other, while the extreme and three extreme method passages show lower comprehension percentages. The differences in comprehension between the passages of the average scoring method showed a statistically significant increase of approximately 11%, $t(17) = -3.47, p = .003$ with a small effect size ($d = -0.82$). There was about a 14% increase from the pre- to the posttest; however, this gain narrowly missed statistical significance under the Holm-Bonferroni procedure, $t(17) = -2.43, p = .026, d = -0.57$. For the passages used in the extreme scoring method, a gain of about 7% was observed, but it was not significant, $t(17) = -1.57, p = .135, d = -0.37$. There was a significant 12% decrease in comprehension for the passages from the three extremes method, $t(17) = -4.43, p < .001, d = -1.04$.

Table 8. *Comprehension percentage means, standard deviations, 95% confidence intervals, and gains of the four scoring methods for the TR group (n = 18)*

Scoring method	<i>M</i>	<i>SD</i>	95% <i>CI</i>
Average			
First 3	61.11	18.96	[51.68, 70.54]
Last 3	72.41	19.36	[62.78, 82.04]
Gains	11.30**	13.83	[4.42, 18.18]
Last minus first			
Pretest	56.94	20.21	[46.89, 67.00]
Posttest	70.83	15.46	[63.15, 78.52]
Gains	13.89*	24.21	[1.85, 25.93]
Extreme			
Slowest	56.94	20.21	[46.89, 67.00]
Fastest	63.89	15.98	[55.94, 71.84]
Gains	6.95	18.80	[-2.40, 16.29]
3 Extremes			
Slowest 3	62.04	15.84	[54.16, 69.92]
Fastest 3	49.72	9.72	[44.89, 54.55]
Gains	-12.32**	11.80	[-18.18, -6.45]

Note. First 3 = first three passages of the treatment. Last 3 = last three passages of the treatment. Gains = increase in comprehension percentage. 1st = first passage of the treatment. 30th = thirtieth passage of the treatment. Slowest = slowest passage of the treatment. Fastest = fastest passage of the treatment. Slowest 3 = slowest three passages of the treatment. Fastest 3 = fastest three passages of the treatment.

* = $p \leq .05$. ** = $p \leq .01$

RQ3 investigates how the reading fluency treatment groups, along with a comparison group, compare with each other in terms of reading rate and reading comprehension by the end of the treatment period. Because the groups were not initially equal in terms of proficiency, a one-way MANCOVA was employed to compare the between-group differences. The independent variable, which was the group, included three levels: the TROR, TR, and comparison group. The dependent variables were the posttest reading rate and posttest comprehension scores. The covariates were the pretest reading rates and pretest comprehension scores.

A preliminary analysis evaluating the homogeneity of regression slopes assumption indicated that the relationship between the covariates and the dependent variable did not differ significantly as a function of the independent variable, as indicated by the non-significant interaction between the pretest reading rate covariate and the independent variable, $F(4, 84) = 0.94, p = .444$, partial $\eta^2 = .04$, as well as the non-significant interaction between the pretest comprehension covariate, $F(4, 84) = 1.08, p = .374$, partial $\eta^2 = .05$. Other preliminary analyses assessing the independence of the covariate and treatment effects were conducted. Non-significant results for both the pretest reading rates, $F(2, 52) = 2.04, p = .141$, and comprehension scores were found, $F(2, 52) = 0.14, p = .872$. These preliminary results indicate that the pretest reading rate and comprehension scores were appropriate to use as covariates in the analysis.

The MANCOVA results indicated that significant differences were found among the three groups on the dependent measures, Wilks's $\Lambda = 0.66$, $F(4, 98) = 5.71$, $p < .001$. The strength of the relationship between the groups and the posttest scores assessed by a partial η^2 showed the group variable accounted for 18.90% of the variance of the dependent variables. The posttest reading rates and comprehension scores adjusted for initial differences are ordered according to their means. The TR group had the largest adjusted mean ($M = 107.98$; 95% $CI = 98.93, 117.04$; $SE = 4.51$), followed by the TROR group ($M = 103.96$; 95% $CI = 95.21, 112.72$; $SE = 4.36$), and the comparison group ($M = 77.58$; 95% $CI = 68.09, 87.06$; $SE = 4.73$). For comprehension, the TROR group had the largest adjusted mean ($M = 70.4$; 95% $CI = 63.40, 77.4$; $SE = 3.50$), followed by the TR group ($M = 68.0$; 95% $CI = 61.30, 74.8$; $SE = 3.40$), and the comparison group ($M = 64.50$; 95% $CI = 57.20, 71.90$; $SE = 3.70$).

One-way ANOVAs on the dependent variables were conducted as follow-up tests to the MANCOVA using the Bonferroni method. For posttest reading rates, a significant difference was found, $F(2, 50) = 12.13$, $p < .001$, partial $\eta^2 = .33$. However, for posttest comprehension scores, the difference was not significant, $F(2, 50) = 0.67$, $p = .514$, partial $\eta^2 = .03$.

Post hoc analyses to the univariate ANOVA for posttest reading rates consisted of conducting pairwise comparisons using the Bonferroni procedure to find how the groups differed from each other in terms of reading rate. The results indicated that the adjusted mean for the TR group differed significantly from the comparison group ($p < .001$) and the TROR group also differed significantly from the comparison group in terms of rate ($p = .001$). There were, however, no significant differences between the treatment groups ($p = 1.00$).

Discussion

RQ1 asked to what extent the treatment groups progressed in their reading rates over one academic semester. According to the average, last minus the first, extreme, and three extremes scoring methods used to assess reading rate gain, the TROR group increased their reading speed by approximately 15, 13, 27, and 21 swpm, respectively, while the TR group made increases of 17, 15, 27, and 23 swpm, respectively. These results are promising because all within-subjects rate gains by both groups were statistically significant with modest but non-trivial effect sizes. Furthermore, the progression throughout the treatment for both experimental groups also indicated upward trends surpassing reading speeds of more than 100 swpm. This result is important because Nation (2005) commented that speeds of less than 100 wpm could hinder comprehension, memory retention, and concentration. In contrast, the comparison group had a significant decrease in reading rate from the pre- to posttest despite having the highest initial mean. This finding suggests that reading abilities might not progress and losses in fluency can occur if reading rate is not explicitly targeted. The evidence also suggests that the reading fluency treatments in this study gave the learners repeated opportunities to practice reading faster with productive pressure, which has facilitated the automatization of word recognition and chunking abilities (DeKeyser, 2007; LaBerge & Samuels, 1974). It also gave these learners the chance to practice reading smoothly from left to right and break stifling L2 reading habits (Takase, 2003).

Comparing the progression of the two treatment groups, the TR group consistently read faster than the TROR group; however, the TROR group nearly caught up to the TR group by the 10th week. The TR group reached higher reading rates, with the highest weekly average being approximately 113 swpm, while the highest weekly average for the TROR group was about 104 swpm. This difference is likely attributed to overall proficiency, as the TR group was slightly more proficient and had a higher initial reading rate. However, there were noticeable regressions in rate in week 5 and 7 for the TR group. The first dip might indicate that the participants were still struggling to increase their reading rate while maintaining sufficient comprehension. The second dip might have occurred because the reading passages doubled in length from passages 16 to 20 and participants experienced some task fatigue. This decrease also suggests there might be an acclimation period for reading longer and more difficult passages where reading rates might regress. Despite the regressions, the TR group participants were able to resume an upward reading rate progression. While regressions in rate were observed for the TROR group as well, they were less dramatic, and might suggest how oral rereading and chunking practice can promote stable increases in reading rate.

The rate gains of the TR and TROR groups are modest in comparison to past studies. While Chung and Nation (2006) reported large gains of 73, 132, and 97 wpm according to the average, highest minus the lowest passages, and the 20th minus first passage, respectively, reading comprehension was not reported, so it remains inconclusive whether or not true fluency was achieved among their participants. Moreover, some reading was done outside the class, which calls into question the reliability of some of the data, and statistical significance testing was not carried out. Information regarding the learners' general English and reading proficiencies were also lacking, making it more difficult to compare the results to the current study. Chung and Nation also reported that most of the gains were observed in the first 10 reading passages. An alternative interpretation to gains in the first 10 texts might be that participants were still acclimating to the reading task instead of developing fluency. Hence, the gains reported might look deceptively larger than in actuality. Tran (2012) also reported greater treatment group gains ranging from 51 to 98 wpm according to same four scoring methods used in the current study. However, again, the learners' proficiency is unclear, making comparisons difficult. A similarity of the current study to Tran's study is the participants in this study showed consistent growth, and gains were seen toward the end of the treatment period.

The gains in the study are more akin to the Taiwanese learners in Chang (2010), where the treatment group increased their reading rate by about 29 wpm (118 to 147 wpm). The timed reading group in Chang (2012) gained about 50 wpm, while the repeated oral reading group gained about 23 wpm, in which the latter gain is similar to the gains of the TR and TROR groups. The learners in Chang and Millet (2013) and Underwood et al. (2012) made rate gains ranging from 40 to 50 wpm by the end of the treatment period. These gains are also comparable to the ones observed in the current study.

Overall, while reading rate gains for both treatment groups were not as large as those reported in prior studies, it should be reiterated that reading rates measured in this study were in standard words per minute. The numerical values for rate in standard words per minute are usually slightly lower than regular words per minute calculations and have been estimated to be approximately 15% lower (Beglar & Hunt, 2014). Moreover, the reading proficiency of the

learners in this study was most likely lower than the learners in the other studies. Also, all participants were continuously encouraged to achieve a minimum of a 70% comprehension rate, even if it meant reducing their reading speed.

RQ2 asked to what extent the treatment groups progressed in their reading comprehension over one academic semester. The results showed that as reading rate increased, comprehension rates for both groups were generally maintained throughout the treatment period, with upward trends toward the end. For the TROR group, there was an upward trend, and the average for weeks 8, 9, and 10 surpassed the 70% comprehension threshold, suggesting that prosody training can enhance comprehension. This result might be due to the fact that the oral reading practice included a phonological component in addition to orthographic and semantic processes involved in silent reading; therefore these learners could benefit from higher quality lexical input and output practice (Perfetti & Hart, 2002). For the TR group, comprehension remained relatively stable throughout the treatment. However, the last part of the treatment period saw a slight decrease. The most likely explanation is that the last 10 passages were from a different series; they were longer, more difficult in terms of readability statistics, and the themes of the passages were more academic. It must be stated, however, that longer passages were intentionally used for the last 10 treatment passages for the TR group in order to minimize differences in time on task and number of words processed for the treatment groups, as the TROR group processed more words with the repeated reading activity despite reading a fewer number of passages.

According to the passages that were used in the four scoring methods, positive gains of 8, 12, and 3% in comprehension were observed for the average, last minus the first, and the three extremes scoring methods, respectively, for the TROR group, and the former two were statistically significant. Moreover, the average of the last three passages for the TROR group surpassed the comprehension threshold, with an average of 72%. The posttest average neared the threshold at 68%. For the TR group, there were increases of 11, 14, and 7% for the average, last minus the first, and extreme scoring methods, respectively, with the former two also being statistically significant. The average of the last three passages for the TR group was 72% and the posttest average was 71%. The comparison group also had a significant increase in comprehension from 54 to 64% on the pre- to posttest, respectively, but this gain was accompanied by a significant decrease in rate. This result suggests that there might be tradeoffs between reading rate and comprehension among L2 learners without reading fluency practice.

When comparing the TROR and TR groups' performance of the same passages, the TROR group started outperforming the TR group in terms of comprehension on passages 14, 15, 16, 18, 19, and 20, and the average of the third phase reached 70%. For the most part, however, the average comprehension for both treatment groups was close but slightly shy of the generally accepted comprehension threshold of 70%. This phenomenon of achieving below 70% comprehension is not uncommon in L2 contexts. Prior reading fluency studies have indicated similar comprehension percentages. For example, the experimental group in Chang (2010) gained 4% on pre- and posttest measures, from 63 to 67%. In Chang's 2012 study, the timed reading group made increases of 14%, from 53 to 67%, and the repeated reading group made an increase of 6%, from 53 to 60%, on the pre- and posttest, respectively. While the repeated reading participants' in Chang and Millet (2013) improved 19% (51 to 70%) and 17% (49 to 66%) for practiced and unpracticed texts, respectively, final comprehension scores were at or slightly below the criterion.

The non-repeated reading students improved by 5% (49 to 54%) and 3% (44 to 47%), respectively, and were well below the 70% benchmark. The Japanese high school students in Underwood et al. (2012) also had comprehension scores around 60%. While these comprehension levels are below the 70% criterion, Carver (2000) has operationalized *rauding* comprehension (the fastest speed where readers can comfortably comprehend a text) at 64% accuracy for L1 learners. Going by this figure, the participants of both the TROR and TR groups fulfilled this comprehension criterion by the end of the treatment period.

RQ3 asked how the groups in the study compare with each other in terms of reading rate and reading comprehension by the end of the treatment period. Comparing pre- and posttest differences between the groups, the TROR and TR groups significantly outperformed the comparison group in terms of reading rate. However, no statistically significant rate differences were found between the two treatment groups, even though the TR group had a slightly higher adjusted mean. Moreover, no statistically significant differences were found between the groups in terms of comprehension, although the TROR group had the highest adjusted mean. These results indicate the effectiveness of timed and repeated oral reading in increasing reading fluency, as opposed to a curriculum that does not specifically target reading fluency development. However, the results also indicate that the superiority of one treatment over the other in this study cannot be stated conclusively. While the TR group ended with a higher mean reading rate compared to the TROR group, the differences were slight. Chang's (2012) timed reading group also showed more comparative gains than the repeated oral reading group. However, the difference was more pronounced. On the other hand, Chang and Millet (2013) found opposite results, as the repeated reading group outperformed the timed reading group on both rate and comprehension measures. In sum, previous literature, as well as the current study have not clearly distinguished the most efficacious reading fluency activity. What is clear is that the learners in the current study benefited from timed reading practice as well as oral reading and chunking practice, which served as a useful scaffolding device that has mutually supported their reading speed and comprehension.

Conclusion

This study has shown the progression of silent reading fluency over one academic semester among lower proficiency Japanese university students. The major findings were: (a) all reading rate gains for the reading fluency groups measured by the average, last minus the first, extreme, and three extreme scoring methods were modest, ranging from 13 to 27 swpm, yet statistically significant; (b) Reading rate gains were accompanied by increases in comprehension except under the extreme method for the TROR group and the three extremes method for the TR group. Comprehension percentages were generally slightly below the 70% threshold throughout the treatment but the progression was trending upward; (c) the treatment groups significantly outperformed the comparison group in terms of rate and had higher comprehension percentages than the comparison group by the end of the treatment period. While no statistically significant differences were found between the treatment groups on rate and comprehension, both treatments were efficacious in promoting reading fluency. It is surmised that timed reading and repeated oral reading provided the productive pressure and necessary practice to enhance the automaticity of word recognition, which has benefited both the reading speed and

comprehension for these learners. The additional oral output and chunking practice that the TROR group received helped their silent reading rate as the participants learned to prioritize content words and read in rhythmic phrases, thereby helping them to read more efficiently. Hence, the findings are optimistic in that even low-proficiency learners can benefit from timed reading and repeated oral reading with chunking practice. It also has been shown that combinations of various fluency activities can be employed as useful variants of regular timed reading practice. In sum, by prioritizing reading faster than usual, drawing attention to prosodic features of English when reading orally, and encouraging the chunking of meaningful phrases, these learners took a significant step to breaking the vicious cycle of slow, painful L2 reading.

This study is not without its limitations. The first limitation is that the length and difficulty of the treatment passages could have been better controlled. That is, the passages ideally should have contained approximately the same number of standard words and the Flesch-Kincaid readability level should have been more similar. Second, the final 10 reading passages for the TR group were noticeably more difficult and had more comprehension questions, which might explain the slight decrease in comprehension toward the end of the treatment period. Another limitation is the proficiency difference between the two treatment groups; however, this imbalance was unavoidable due to curriculum constraints. Moreover, the members of the comparison group were from a different department; therefore, the overall time studying English compared to the treatment groups was not equal. Finally, in order to control for time on task, the number of words processed during the treatment for the TROR and TR groups were not exactly equal.

Future L2 reading fluency studies can be improved in four ways. First, it will be beneficial to examine the progression of reading rates and comprehension for longer than one academic semester. Second, including a variety of participants such as higher proficiency learners, learners of different ages, as well as learners from different language and cultural backgrounds, is needed to gain a better picture of how L2 reading fluency develops across proficiency, age group, and learner backgrounds. Third, study designs and measurement practices should be strengthened. There needs to be more control over the number of words read as well as time on task. Additionally, standard word measurements should be used when assessing reading rate. Larger sample sizes of participants are also needed. Finally, it would be insightful to continue to compare various reading fluency activities. Research comparing different combinations and testing for their differential effectiveness is necessary in order to get a nuanced depiction of how to best develop reading fluency for L2 learners. Empirical evidence pertaining to these issues will bolster the current L2 reading fluency research.

Acknowledgements

I would like to extend my gratitude to Dr. David Beglar of Temple University for all his assistance, expertise, and encouragement during this research project, as well as revising earlier drafts of this manuscript. I would also like to thank the two anonymous reviewers for all their helpful comments. Additionally, I want to thank Michael Holsworth of Kyoto Sangyo University and Christopher Edelman of Ritsumeikan University for helping collect some of the data for this study. Finally, I am grateful for Drs. Paul Leeming and Paul Joyce of Kindai University for their help revising earlier versions of this paper.

References

- Beglar, D., Hunt, A., & Kite, Y. (2012). The effect of pleasure reading on Japanese university EFL learners' reading rates. *Language Learning*, 62, 665–703. doi:10.1111/j.1467-9922.2011.00651.x
- Beglar, D., & Hunt, A. (2014). Pleasure reading and reading rate gains. *Reading in a Foreign Language*, 26, 29–48.
- Breznitz, Z. (2006). *Fluency in reading*. Mahwah, NJ: Erlbaum.
- Browne, C. (1998). Japanese high school textbooks: How readable are they? *Temple University Japan Working Papers in Applied Linguistics*, 8, 28–41. Retrieved from http://www.wordengine.jp/research/pdf/Japanese_high_school_textbook.pdf
- Carver, R. (1990). *Reading rate: A review of research and theory*. San Diego, CA: Academic Press.
- Carver, R. (2000). *The causes of high and low reading achievement*. Mahwah, NJ: Erlbaum.
- Chang, A. C-S. (2010). The effect of a timed reading activity on EFL learners: Speed, comprehension, and perceptions. *Reading in a Foreign Language*, 22, 43–62.
- Chang, A. C-S. (2012). Improving reading rate activities for EFL students: Timed reading and repeated oral reading. *Reading in a Foreign Language*, 24, 56–83.
- Chang, A. C-S., & Millet, S. (2013). Improving reading rates and comprehension through timed repeated reading. *Reading in a Foreign Language*, 25, 126–148.
- Chung, M., & Nation, P. (2006). The effect of a speed reading course. *English Teaching*, 61(4), 181–204.
- Chung, M. (2010). The effect of a speed reading course: A replication. *Asian Journal of English Language Teaching*, 20, 95–116.
- DeKeyser, R. (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. New York, NY: Cambridge University Press.
- Educational Testing Service. (2006). *TOEIC Bridge and TOEIC score comparisons*. Retrieved from https://www.ets.org/Media/Tests/TOEIC/pdf/ToEICBridge_Cmprsn.pdf
- Extensive Reading Foundation. (2011). *The Extensive Reading Foundation's guide to extensive reading*. Retrieved from http://erfoundation.org/ERF_Guide.pdf
- Gibson, S. (2008). Reading aloud: A useful learning tool? *English Language Teaching Journal*, 62, 29–36. doi:10.1093/elt/ccm075
- Gorsuch, G. J. (1998). Yakudoku EFL instruction in two Japanese high school classrooms: An exploratory study. *JALT Journal*, 20, 6–32. Retrieved from <http://jalt-publications.org/jj/articles/2777-yakudoku-efl-instruction-two-japanese-high-school-classrooms-exploratory-study>
- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. New York, NY: Cambridge University Press.
- Grabe, W. (2010). Fluency in reading: Thirty-five years later. *Reading in a Foreign Language*, 22, 71–83.
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6, 65–70.
- LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 6, 293–323. doi:10.1016/0010-0285(74)90015-2
- Macalister, J. (2008). Effect of a speed reading course in an English as a second language environment. *TESOLANZ Journal*, 16, 23–32.

- Macalister, J. (2010). Speed reading courses and their effect on reading authentic texts: A preliminary investigation. *Reading in a Foreign Language*, 22, 104–116.
- Mikulecky, B., & Jeffries, L. (2005). *Reading power: Reading for pleasure, comprehension skills, thinking skills, reading faster* (3rd ed.). New York, NY: Pearson Education.
- Nation, I. S. P. (2005). Reading faster. *PASAA*, 36, 21–35.
- Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31(7), 9–13.
- Nation, I. S. P., & Malarcher, C. (2007a). *Reading for speed and fluency 1*. Seoul: Compass.
- Nation, I. S. P., & Malarcher, C. (2007b). *Reading for speed and fluency 2*. Seoul: Compass.
- Nation, I. S. P., & Malarcher, C. (2007c). *Reading for speed and fluency 3*. Seoul: Compass.
- National Reading Panel. (2000). *Report of the subgroups: National reading panel*. Washington, D.C.: National Institute of Child Health and Development.
- Newfields, T. (2006). Suggested answers for assessment self-study quiz #1. *Shiken: JALT Testing & Evaluation SIG Newsletter*, 10(2), 25–32.
- Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy: Studies in written language and literacy 11* (pp. 189–213). Philadelphia, PA: Benjamins. doi:10.1075/swll.11.14per
- Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, 64, 878–912. doi: 10.1111/lang.12079
- Quinn, E., & Nation, P. (1974). *Speed reading*. Wellington, New Zealand: Victoria University Bookcentre.
- Quinn, E., Nation, P., & Millet, S. (2007). *Asian and Pacific Speed Readings for ESL Learners*. Wellington: English Language Institute Occasional Publication. Retrieved from <https://www.victoria.ac.nz/lals/about/staff/publications/paul-nation/Speed-reading-whole.pdf>
- Rashotte, C. A., & Torgesen, J. K. (1985). Repeated reading and reading fluency in learning disabled children. *Reading Research Quarterly*, 20, 180–188. doi:10.1598/rrq.20.2.4
- Rasinski, T., Rikli, A., & Johnston, S. (2009). Reading fluency: More than automaticity? More than a concern for primary grades? *Literacy Research and Instruction*, 48, 350–361. doi: 10.1080/19388070802468715
- Samuels, S. J. (1979). The method of repeated readings. *The Reading Teacher*, 32, 403–408.
- Schrauben, J. E. (2010). Prosody's contribution to fluency: An examination of the theory of automatic information processing. *Reading Psychology*, 31, 82–92. doi:10.1080/02702710902753996
- 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, CA: Academic Press.
- Segalowitz, N. (2003). Automaticity and second languages. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 382–408). Malden, MA: Blackwell.
- Taguchi, E., Gorsuch, G., Lems, K., & Rosszell, R. (2016). Scaffolding in L2 reading: How repetition and an auditory model help readers. *Reading in a Foreign Language*, 28, 101–117.
- Taguchi, E., Gorsuch, G., Takayasu-Maass, M., & Snipp, K. (2012). Assisted repeated reading with an advanced-level Japanese EFL reader: A longitudinal diary study. *Reading in a Foreign Language*, 24, 30–55.
- Takase, A. (2003). Effects of eliminating some demotivating factors in reading English

- extensively. *JALT 2003 Conference Proceedings*, 95–103.
- Takeuchi, O. (2003). *Searching for better language learning strategies: Studies on good foreign language learners in the Japanese FL context*. Tokyo: Shohakusha. Retrieved from <http://www2.ipcku.kansai-u.ac.jp/~takeuchi/papers/Searching.pdf>
- Tran, T. N. Y. (2012). The effects of a speed reading course and speed transfer to other types of texts. *RELC Journal*, 43, 23–37. doi:10.1177/0033688212439996
- Underwood, P., Myskow, G., & Hattori, T. (2012). The effect of speed reading instruction on Japanese high school students' English reading comprehension and vocabulary development. *Journal of International Education Research*, 8, 27–40. Retrieved from https://www.researchgate.net/publication/298911696_The_Effect_Of_Speed_Reading_Instruction_On_Japanese_High_School_Students_English_Reading_Comprehension_And_Vocabulary_Development
- W3Techs – World wide web technology surveys: Content languages. (2017). Retrieved from <https://w3techs.com/>
- Yamashita, J., & Ichikawa, S. (2010). Examining reading fluency in a foreign language: Effects of text segmentation on L2 readers. *Reading in a Foreign Language*, 22, 263–283.

Appendix

Example Chunking Passage for the TROR Group

Passage 1: Susan and Sam

Susan Conley Diamond and Sam Diamond / live in Rosebud, / a small town in New Jersey. / It looks like / many other towns / in the United States. / On Main Street, / there is a post office / and a police station. / The drugstore and the library / are down the street. / There's also a shopping center, / with a supermarket / and a department store.

In the middle of Rosebud, / near the post office, / is Dr. Sam Diamond's office. / Everybody in town / knows Dr. Diamond. / He's a good dentist. / He's also / a popular person. / He likes to tell funny stories / to his patients. / They forget about their teeth / when they listen to him.

Susan Conley is / Sam Diamond's wife. / She's a scientist / with a Ph.D. / in biology. / She works with / a group of scientists / in a laboratory / in New York City. / They're studying / the human brain / and looking for ways / to help people / with Alzheimer's / and other serious diseases.

Susan usually takes the train / from Rosebud to New York. / Sometimes she stays at home / and works on her computer. / She's very happy / when she can work at home. / But she likes working / in the lab / with interesting people, / and she likes being / in an exciting place / like New York.

About the Author

Torrin R. Shimono is a lecturer in the Law Department at Kindai University and a part-time instructor in the Business Administration Department at Ritsumeikan University. He has an M.S.Ed. from Temple University in TESOL and is currently a Ph.D. candidate in the seventh doctoral cohort at Temple University, Japan for applied linguistics. His research interests include

L2 reading and speaking fluency, reading self-efficacy, pronunciation, and motivation. Email: torrin@jus.kindai.ac.jp