



All Theses and Dissertations

2015-07-01

Improving Reading Fluency and Comprehension in Intermediate L2 Japanese Learners

Traci Tomi Andreason
Brigham Young University - Provo

Follow this and additional works at: <https://scholarsarchive.byu.edu/etd>



Part of the [Other Languages, Societies, and Cultures Commons](#)

BYU ScholarsArchive Citation

Andreason, Traci Tomi, "Improving Reading Fluency and Comprehension in Intermediate L2 Japanese Learners" (2015). *All Theses and Dissertations*. 5480.

<https://scholarsarchive.byu.edu/etd/5480>

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in All Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.

Improving Reading Fluency and Comprehension in
Intermediate L2 Japanese Learners

Traci Tomi Andreason

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Arts

J Paul Warnick, Chair
Robert Erickson
Masakazu Watabe

Center for Language Studies

Brigham Young University

July 2015

Copyright ©2015 Traci Tomi Andreason

All Rights Reserved

ABSTRACT

Improving Reading Fluency and Comprehension in Intermediate L2 Japanese Learners

Traci Tomi Andreason
Center for Language Studies, BYU
Master of Arts

This study explores the effects of different methods employed by language teachers, namely assisted repeated reading (Taguchi, Gorsuch, Takayasu-Maass, & Snipp, 2012) and morphological and syntactic explanation (Tong, Deacon, & Cain, 2013), to improve language learners' reading fluency and reading comprehension, respectively. Volunteers from students enrolled in Japanese 301 through 322 classes at Brigham Young University (BYU) participated in this study. The participants (a) submitted a background questionnaire and consent form, (b) attended a one-hour session in the testing lab, and (c) attended a second one-hour session in the testing lab. Each testing session consisted of the participants making a pre- and post- reading recording and taking reading comprehension tests for two separate reading scripts. Each participant submitted a total of eight reading recordings and eight sets of answers to five reading comprehension questions. The reading texts were chosen from 3-kyu (3rd level of 5, which was deemed to be intermediate by the researcher) of the Japanese Language Proficiency test (JLPT), with a minor modification. Each recording and answer sheet submission was then graded and recorded. The differences between the pre- and post-test scores were then examined to determine the effectiveness of each treatment employed. The treatments included were: (a) assisted repeated reading, (b) assisted repeated reading followed by morphological and syntactic explanation, (c) morphological and syntactic explanation, and (d) morphological and syntactic explanation followed by assisted repeated reading. The analysis of the resultant data revealed that although there was a statistically significant increase in scores for each of the methods and there were differences in the scores between the different methods, there was not a significant difference between the methods. However, one of the important implications that can be drawn from this study is that a short fluency-building exercise of 2-5 minutes can yield significant gains in the language learners, in both reading fluency and in reading comprehension.

Keywords: reading fluency, reading comprehension, Japanese, repeated reading, assisted repeated reading, morphological and syntactic explanation, language learning, language teaching

ACKNOWLEDGEMENTS

I would like to thank all those who have helped me in my academic life. My parents, Theodore H. and Shizuka S. Okawa, have given me their love and support through the most trying times in my life. My husband, Adam J. Andreason, has supported me emotionally and academically. I would also like to express my gratitude for the following individuals: my thesis committee chair, Dr. Paul Warnick, for his excellent example and being a wonderful mentor, Dr. Robert Erickson and Dr. Masakazu Watabe for their help and guidance, Dr. Russell A. Hansen and Mr. Devin N. Asay of the Office of Digital Humanities for their help in the implementation of the research, and Dr. Dennis Eggett for his help in the statistical analysis of my research data.

Table of Contents

Abstract.....	ii
ACKNOWLEDGEMENTS	iii
Table of Contents	iv
List of Tables	vi
Chapter 1: Introduction	1
<i>Introduction.....</i>	<i>1</i>
<i>Research Questions.....</i>	<i>2</i>
Chapter 2: Review of Literature	4
<i>Reading Fluency.....</i>	<i>4</i>
<i>Relationship between reading fluency and reading comprehension</i>	<i>5</i>
<i>Methods for gaining reading fluency</i>	<i>7</i>
<i>Methods for gaining reading comprehension.....</i>	<i>9</i>
<i>Unique challenges of reading in Japanese.....</i>	<i>10</i>
<i>Methods of measuring reading fluency and comprehension</i>	<i>11</i>
<i>Summary.....</i>	<i>13</i>
Chapter 3: Methodology	15
<i>Participants.....</i>	<i>15</i>
<i>Timetable.....</i>	<i>17</i>
<i>Instruments.....</i>	<i>18</i>
<i>Procedure.....</i>	<i>20</i>
<i>Grading.....</i>	<i>25</i>
Chapter 4: Results and Discussions	28
<i>Results and Data Analysis.....</i>	<i>28</i>
Chapter 5: Summary and Conclusion	41
<i>Summary.....</i>	<i>41</i>
<i>Implications.....</i>	<i>42</i>
<i>Limitations of the Study.....</i>	<i>43</i>
<i>Suggestions for Future Research</i>	<i>45</i>

References.....	47
Appendices.....	50
<i>Appendix A</i>	50
<i>Appendix B</i>	51
<i>Appendix C</i>	53
<i>Appendix D</i>	54
<i>Appendix E</i>	58
<i>Appendix F</i>	66
<i>Appendix G</i>	67

List of Tables

Table 1 Participants.....	17
Table 2 Timetable.....	18
Table 3 Test Sessions	21
Table 4 One Hour Estimation of Testing Sessions.....	23
Table 5 Descriptive statistics of group A	29
Table 6 Descriptive statistics of group B	30
Table 7 Analysis of RF Accuracy	31
Table 8 Analysis of RF Speed.....	31
Table 9 Analysis of Reading Comprehension.....	32
Table 10 Inter-rater Reliability data	37
Table 11 Inter-rater Reliability Correlation.....	37
Table 12 Differences between Treatments in Analysis of RF accuracy.....	38
Table 13 Differences between Treatments in Analysis of RF Speed	39
Table 14 Differences between Treatments in Analysis of Reading Comprehension	38
Table 15 Gains Observed in Each Treatment.....	40

Chapter 1: Introduction

Introduction

Much research has been done on how language learners can improve their reading comprehension (LaBerge & Samuels, 1974; Tong, Deacon, & Cain, 2013) and how they might improve their reading fluency (Kim, Wagner, & Lopez, 2012; Ocana, 2011; Schmidt, 1992; Taguchi, Gorsuch, Takayasu-Maass, & Snipp, 2012). There have also been studies on the correlation between the two (Shen & Jiang, 2013; Kim, Wagner, & Lopez, 2012; Jodai & Tahriri, 2011). However, as in many aspects of research, there have not been as many second language (L2) studies done on these subjects as there have been first language (L1) studies (Shen & Jiang, 2013). There have been even fewer studies on Truly Foreign Languages (TFLs) in these subjects. TFLs, as described by Ning (2009), are defined as languages with few cognates and commonalities in “their grammar, vocabulary, and writing systems” (p.29). For that reason, languages such as Chinese, Japanese, Korean, and Arabic are often referred to as “truly foreign” (p.29) to English speakers. The only similar study related to an Asian TFL is that of Shen and Jiang (2013). Their study researched reading fluency and reading comprehension using Chinese learners as participants.

Though the languages that fall into the category of TFL seem to receive little attention in the literature in general, because of their very nature of being truly foreign they especially need research in order to help foster more successful language learners. According to the American Council on the Teaching of Foreign Languages (ACTFL), it is estimated that a learner of average aptitude could increase their proficiency to the “Advanced low” level with 720 hours of classroom instruction in languages such as Spanish, French, and Romanian. By contrast, it would

take 1320 hours to achieve the same proficiency for average learners of TFLs (Language Testing International, 2014). That is almost double the classroom instruction time.

Additionally, because of the foreignness of their orthography, TFL learners often have a separate struggle with reading proficiency from that for oral proficiency. Jorden and Walton (1987) explain the challenges of learning a TFL thus: “Those TFLs that use nonalphabetic scripts create an entirely new learning task for the student. The problems begin at the graphic level, but extend in some cases—as with Chinese characters—to the level of forming a sound-symbol relationship where the symbol may give no clue to the sound and vice versa” (p.117).

Proficiency, in the scope of this study, will be measured according to the ACTFL proficiency level and estimated classroom instruction time as cited above. The participants in this study, though TFL learners, are estimated to be intermediate Japanese learners.

Research Questions

This study explores the following research questions:

1. How effective is Assisted Repeated Reading, where the learners repeatedly practice reading the same passage in the target language while having access to a native speaker’s audio example, in improving reading fluency for intermediate Japanese learners? In this case, intermediate Japanese learners will refer to students that have or are currently enrolled in Japanese 301 or higher where the learners have studied all 92 of the basic phonetic symbols and over 400 of the Chinese characters, called kanji.
2. How effective is morphological and syntactic explanation in improving reading comprehension for intermediate Japanese learners? The reading comprehension will be measured by a pre- and post- reading comprehension test consisting of 5 questions.

3. Will Assisted Repeated Reading also influence reading comprehension and will morphological and syntactic explanation also improve reading fluency? Reading fluency will be measured in two parts, through accuracy in reading characters and through speed. The study's focus will be on the differences between the learners' pre-test and post-test scores of their reading comprehension and reading fluency tests.
4. When the methods are employed in a different order, will there be a difference in the students' level of improvement in their reading fluency and reading comprehension? This can have important pedagogical implications for language learning especially in lesson planning.

The independent variables for this study are the two different treatments, assisted repeated reading (ARR) and morphological and syntactic explanation (explanation), and their two different combinations, ARR-explanation and explanation-ARR. The dependent variables are reading fluency and reading comprehension. Reading fluency is measured in accuracy and speed. The accuracy is measured as a percentage of the number of correct characters that were read divided by the total number of characters that the participant attempted to read in the reading script. The speed is measured in the number of characters read, whether correct or incorrect, divided by the time it took to read them.

Chapter 2: Review of Literature

Reading Fluency

Schmidt (1992) describes fluency as “the processing of language in real time” (p.358). The process can apply to any of the four skills (i.e. reading, writing, speaking, and listening) but would be an internal processing—something the student must do while utilizing their interlanguage. According to Lightbown & Spada (2013), the interlanguage is “a learner’s developing second language knowledge” (p.220) and may have characteristics of their first language but is dynamic and therefore will continue to be revised and be added upon as they learn the language. When a learner can process more readily the target language, the learner is said to have become more fluent in that language.

Nation and Newton (2009) name fluency as one of the four essential strands in language education. The authors state:

A well-balanced language course should consist of four roughly equal strands: 1. Learning through meaning-focused input . . . 2. Learning through meaning-focused output . . . 3. Learning through deliberate attention to language items and language features . . . 4. Developing fluent use of known language items and features over the four skills of listening, speaking, reading, and writing; that is, becoming fluent with what is already known. (pp. 1-2)

Nation and Newton (2009) further recommend that the four strands should receive equal attention in the classroom. According to this recommendation, a quarter of the class time should be spent toward bettering the students’ fluency in one or more of the four language skills of reading, listening, writing, and speaking. However, fluency is an often-neglected aspect of the language classroom (Rossiter, Derwing, Manimtim, & Thomson, 2010). Rossiter, Derwing,

Manimtim, and Thomson (2010) further state that textbooks and resource materials are “sorely lacking in consciousness-raising activities” (p. 583) that help students’ automaticity in the language and thereby better their fluency in the language skills. The authors introduce types of oral fluency instructions that can be incorporated into L2 classes to “address these deficiencies” (p.583). These authors, along with Nation and Newton (2009), recommend that whatever the language skill may be that is receiving focus in class, it ought to be accompanied with fluency building activities.

Although there is a rising interest in increasing oral and reading fluency in students (Murray, Munger, & Clonan, 2012), this advancement in fluency should not be done at the expense of the students’ comprehension. Reading smoothly, with accuracy and speed, without understanding the text benefits the student very little in real life. Fortunately, reading fluency and comprehension may be enhanced in the students concurrently. As cited above, Schmidt (1992) attributes fluency advancement to the students’ developing interlanguage. He cites and compares many theories as he reflects on his belief that one cannot improve in fluency without also improving one’s understanding of the language.

Relationship between reading fluency and reading comprehension

When considering fluency, one must also consider the content of the language material and the accuracy of production or language output (e.g. reading when read aloud), since simpler content would naturally be easier to read and comprehend. Nation and Newton (2009) define fluency, accuracy, and complexity of content in the following statement: “Fluency is typically measured by speed of access or production and by the number of hesitations; accuracy by the amount of error; and complexity by the presence of more complicated constructions, such as subordinate clauses” (p. 152).

Kim, Wagner, and Lopez (2012) conducted a study specifically examining the relationship between reading fluency and reading comprehension with L1 English-learning 1st and 2nd grade elementary school children. They explain:

With further development of reading skills . . . children's text-reading fluency may become differentiated from text-free word reading (i.e., list reading fluency). The rationale is that when decoding of individual words develops sufficiently, other factors such as oral language proficiency exert more of an influence on text reading, resulting in text reading fluency becoming uniquely related to reading comprehension . . . (p. 94)

The above quote from Kim, Wagner, and Lopez (2012) suggests several relationships seen between reading fluency and reading comprehension. One of the most important inferences is that contextual reading fluency (i.e. reading in context) influences reading comprehension. This seems a reasonable deduction since language learners are better able to decode individual words more proficiently as they become more fluent readers.

The opposite relationship has also been suggested by Fox (2006). She says that greater reading fluency will help the information move more quickly into long-term memory and therefore will help the reader comprehend better. She states the following:

Reading rate affects comprehension. Information enters short-term memory before it is moved to long-term memory, where the reader stores ideas and makes sense of text. Short-term memory holds only a small amount of information and the information stays in short-term memory only a brief period of time. When the pace of reading is too slow, the reader does not move information quickly into long-term memory. Information in short-term memory creates a roadblock that prevents new information from entering, and

so less information is moved on to long-term memory. Consequently, plodding readers do not grasp as many ideas as fluent readers. (p. 1)

Though they differ between reading fluency and comprehension on which component influences the other, these studies suggest that there is a connection between the two. Although it may not be possible to conclude that an increase in one is the cause of an increase in the other, these findings suggest an observable relationship between the two. Though a causal relationship may still be difficult to prove, the answer to research question #3 should provide some support for the existence of a relationship between reading fluency and reading comprehension.

Methods for gaining reading fluency

There are several ways of practicing reading so as to improve reading fluency. Taguchi, Gorsuch, Takayasu-Maass, and Snipp (2012) use a method called Assisted Repeated Reading in their study. They explain that repeated reading was a method developed by S.J. Samuels as a pedagogical approach for developing English L1 readers' fluency. Samuels' Automaticity Theory explained how automaticity or automatic performance developed in language learners (LaBerge & Samuels, 1974). This Repeated Reading (RR) method has been a popular topic for studies since its development in 1974 and has been tried in different variations such as with or without assistance as in Taguchi, Gorsuch, Takayasu-Maass, and Snipp (2012). Taguchi, Gorsuch, Takayasu-Maass, and Snipp (2012) further explain that RR has undergone extensive research and conclusively shows positive results for improving reading fluency and comprehension in the English L1 setting. However, a strong correlation between improved reading fluency and enhanced comprehension in L2 fluency has yet to be shown (pp. 2-12, 30). In their study, they employed this RR method with assistance, where the learner practiced reading the same passage repeatedly and had access to a native reader's reading to use as an example.

Taguchi, Gorsuch, Takayasu-Maass, and Snipp (2012) conducted a study with one participant to investigate the effects of the RR treatment. The participant was a 34-year old Japanese housewife at an advanced level of English reading ability. The subject completed 70 RR treatments over a 14-week period. The procedure for the study was as follows: the subject read a passage from the RR text silently while timing her own first reading with a stopwatch; read the same passage three more times while listening to the audio recording; read the same passage two more times silently while timing each reading with a stopwatch; and wrote thoughts and comments about the RR session in a diary. The authors then analyzed the subject's diary and her perception, along with the recorded reading speeds, on how her reading fluency and comprehension have improved during this process. While noting that their study is largely subjective by design, the authors concluded that though the RR seems to provide scaffolding in helping readers' vocabulary, the subject commented that repeated reading of a passage with too many unknown vocabulary words or too many unknown grammatical items did not enhance her comprehension of that language material and therefore RR did not seem to improve reading comprehension for new materials. The participant and researchers did not specify how much unknown items were too many, but Hu and Nation (2000) estimated that the text must contain at least 95% of known items for the reader to be able to infer the meaning of the unknown items.

Though the study by Taguchi, Gorsuch, Takayasu-Maass, and Snipp (2012) is a single-subject study and not very generalizable, her comments and the findings of the study have various implications for future research with RR in L2 settings, especially in a Truly Foreign Language (TFL) such as Japanese. Though some researchers believe that greater fluency will improve comprehension as well, it could be seen from this study that RR is helpful in improving reading fluency and in helping with recall of known lexical and grammatical items but may not

help with previously unknown items. More simply, RR helps to improve reading fluency but only improves reading comprehension in as much as most of the grammar or vocabulary in the text have already been introduced to the reader. If this is true, assisted repeated reading may not help the learners to have significant gains in their reading comprehension scores. This is the hypothesis behind research question #3.

Nation & Newton (2009) outline the guidelines for constructing reading fluency lessons and activities as follows: (a) the activity should be meaning-focused, (b) the language items in the activities are within the previous experience of the students, and (c) students receive support and encouragement to perform at a higher than normal level. The second point that these authors mention, that the language items in the activity should be mostly within the students' previous experience, echoes the findings from Taguchi, Gorsuch, Takayasu-Maass, and Snipp's (2012). Nation and Newton (2009) recommend that these three guidelines should be followed for all fluency-building activities, including activities for building reading fluency.

Methods for gaining reading comprehension

Tong, Deacon, and Cain (2013) conducted a study with L1 English students, who were 4th grade elementary school children, with poor reading comprehension abilities in L1. They explain that "research to date indicates that poor reading comprehension is associated with weak metalinguistic skills but that these weaknesses are not universal" (p. 24). In their study, they set out to show that poor reading comprehension is actually associated with lack of morphological and syntactic awareness. The study was conducted with 30 students, 15 with average reading comprehension abilities and 15 that performed poorly on reading comprehension. These 30 participants performed tasks (tests) in the following categories (in order): vocabulary, word reading and passage comprehension, phonological awareness, morphological awareness word

analogy, syntactic awareness sentence correction, morphological awareness sentence completion, and reading comprehension. The result suggested that “. . . poor comprehenders have difficulties with both morphological awareness and syntactic awareness but not phonological awareness” (Tong, Deacon, & Cain, 2013, p. 29).

The findings from Tong, Deacon, and Cain (2013) suggest that lessons with more focus on morphological and syntactic features would help readers to have better reading comprehension. On the other hand, Ocana (2011) conducted a study on the effects of extended explicit systematic phonics instruction on adult L2 reading fluency and was not able to produce conclusive results that phonics or phonological instruction improves reading fluency. Ocana’s (2011) result seems to support the conclusion for Tong, Deacon, & Cain’s (2013), since they also concluded that the focus should be in raising the students’ morphological and syntactic awareness rather than their phonological awareness.

Unique challenges of reading in Japanese

As mentioned above, TFLs have very foreign writing systems, grammar, and vocabulary for L1 English speakers. Because of this foreignness of the language, there are some unique challenges for language learners in reading TFLs. Japanese, as one of the TFLs for native English speakers, also has a very unique writing system that offers additional challenges for Japanese learners. First of all, Japanese has three writing systems. Sakade (2003) called the phonetic alphabet systems “syllabaries (hiragana and katakana)” (p.v). There are 92 of these symbols in the two syllabaries that were formally approved by the government in 1951. Aside from the phonetic syllabaries, there are “1,945 Chinese characters contained in the officially approved character list—the Jōyō kanji (General Use Characters) List” (Sakade, 2003, p.v). While the syllabaries are comparatively short lists of finite sounds, the Chinese characters called

kanji offer a more substantial challenge. There are different readings possible for each kanji, and with different combinations of kanji, a reader may encounter different combinations of sounds.

The explanation Sakade (2003) offers on Japanese orthography depicts this challenge:

Typically, each Chinese character has two types of readings—on-yomi and kun-yomi . . .

In some cases, a given Chinese character has several *on* readings, reflecting different forms of underlying Chinese pronunciation. A given character may also have more than one associated *kun* reading. Context and the use or absence of accompanying kana (okurigana) are the pointers as to which reading is appropriate in a given case. (p.vi)

For beginning to intermediate Japanese learners with limited context in the language and culture, deciphering when a character takes a certain reading can be very challenging. However, as Sakade (2003) cited above, this understanding of context is essential in knowing the reading of the kanji.

The fact that learning to read Japanese has unique challenges may be a factor that may produce contradictory results to those of the studies from Ocana (2011) and Tong, Deacon, & Cain (2013). It may be true that, in general, the focus should be in raising the students' morphological and syntactic awareness rather than their phonological awareness to improve reading fluency. However, Japanese learners may understand the meaning of the word and the syntax, but be unable to produce the phonological reading of given characters that are appropriate to that context.

Methods of measuring reading fluency and comprehension

There have been several measurements of reading fluency used in various studies. Ocana (2011) utilized a holistic rating of recorded reading (i.e. the raters judged the participants' reading as a whole to be correct or incorrect instead of analyzing factors or scales of correctness)

in a study with college-aged ESL students. The grading was executed by two separate raters and Ocana (2011) accounted for inter-rater reliability. In her study, the rater decided if the participants read the word on the screen accurately or not and recorded the ratings on the computer by pressing either Y for yes or N for no.

Shen and Jiang (2013) conducted a study on reading fluency, word segmentation accuracy, and reading comprehension in L2 Chinese with college students. They measured their subjects' reading fluency and reading comprehension with the following methods: First, they measured the students' reading fluency through a process that included the subjects' character-naming speed, their score on a word segmentation test, and a computation of their word segmentation error rate ($\#$ of total word segmentation errors / $\#$ of total characters covered for the word segmentation). Second, they measured the students' comprehension by having them take a comprehension test on the text that was read. Both methods of measurements were deemed to be valid and reliable by the researchers after correlating the scores for forms A and B that were used for each test. The correlation showed that the test reliability was very high with $r=0.92$ ($p<.01$) for character-naming speed, $r=0.91$ ($p<.01$) for character-naming accuracy, and $r=0.97$ ($p<.01$) for the reading comprehension test.

Ocana (2011)'s study showed that, in the first semester, students with phonics instruction were able to decode words better than the control group that did not receive phonics instruction, but failed to show a statistically significant difference between the two groups in the second semester of study. Shen and Jiang (2013) showed that the students that were immersed in the language and culture by studying Chinese in China showed better performance overall: "an increase of 16% for character-naming accuracy, an increase of 17% in the character-naming speed, and a 29% decrease in the word segmentation error rate . . . an increase in reading

comprehension rate by 14.8%” (p. 15). They measured reading fluency in accuracy, speed, and prosody and measured reading comprehension by use of reading comprehension questions.

Jeon (2012) conducted a study on the role of oral reading fluency (ORF) in second language reading in South Korea. In her study, she assessed 255 high school students in oral reading fluency through a total of nine tests including three oral reading fluency tests. The three oral reading fluency tests were a pseudoword reading test, a word reading test, and a passage reading test. In particular, the passage reading test was used to measure “passage reading fluency, which was defined as the ability to orally read a connected text fast and accurately” (p. 191). Unlike Shen and Jiang (2013), Jeon (2012) decided not to include prosody in her study because, after researching several previous studies conducted on similar topics, she concluded that “it was difficult to achieve an acceptable reliability with their prosody measures due to the subjective nature of judging desirable prosody, and that prosody failed to explain a significant amount of reading variance” (p. 191). In the end, Jeon (2012) concluded that oral passage reading fluency alone accounted for 20.9% of variance in silent reading comprehension.

Summary

As Ocana (2011) and Shen and Jiang (2013) mention in the conclusions of their respective studies, more research is needed for better teaching strategies for reading fluency with conclusive results. Compared to studies on L1 English reading comprehension, there has been much less work done on L2 reading comprehension, and even less for L2 Japanese. Though Shen and Jiang (2013) is one of the few that studied the correlation between reading comprehension and reading fluency in a TFL, because of the design of the study (i.e. a lack of comparable conditions between the test groups such as different textbooks, different professors, and different

demands on vocabulary learning), they were unable to produce conclusive results from their findings.

Multiple authors have focused their attention on attaining greater reading comprehension while improving the learners' understanding of the language and productive output, also known as automaticity (Schmidt, 1992; Shen & Jiang, 2013). Most teachers would agree that producing students who can read fluently with accuracy but with little to no comprehension of what they read, is not the ideal result. While many researchers place importance on fluency with comprehension, there have not been many studies on how to improve both fluency and comprehension in Asian languages.

This research will: (a) investigate whether teaching methods for increasing reading fluency (specifically the assisted repeated reading method) are effective, (b) investigate whether teaching methods for reading comprehension (morphological and syntactic explanation) are effective, (c) investigate whether assisted repeated reading could also influence reading comprehension and whether morphological and syntactic explanation could also influence reading fluency; and (d) investigate, when both are given focus, whether the order of application of the two methods will make a significant difference (i.e. comprehension first, then fluency, or vice versa).

Chapter 3: Methodology

Participants

Volunteers were solicited from students enrolled in Japanese 301, 302, 321, or 322 at Brigham Young University, Provo, during winter semester 2015. The volunteers were not monetarily or otherwise compensated. However, the instructors for their respective classes had agreed to give each participant partial or full extra credit for partial or full participation in the study, respectively.

The participants were male and female between the ages of 18 and 30, from various backgrounds. In order to distinguish between heritage learners (those who had and grew up with one or more native Japanese speakers as a parent) and returned missionaries (those who served an 18- to 24-month volunteer mission for the Church of Jesus Christ of Latter-day Saints in Japan) from those whose learning has been mostly in class, each participant was asked to submit a background questionnaire (see Table 1).

The volunteers were then sorted into two groups (A or B) and informed via e-mail to take their corresponding tests during preset time intervals. The group sorting took place using a stratified randomization of the volunteers. The strata were made on two levels: what class they were enrolled in, and then whether or not they were returned missionaries and/or heritage learners. Each stratum was then sorted randomly into either group A or B. Specifically, the stratum of heritage learners from Japanese 301 were sorted into two groups (group A or B), then the returned missionaries from Japanese 301 were sorted into two groups (group A or B), and the rest into two groups (group A or B), and so on for all of the classes. For the volunteers that were both a returned missionary and a heritage learner, they were put in the strata of heritage learners,

rather than returned missionaries to distinguish them from those who were only returned missionaries.

In the beginning, there were 122 volunteers. These were sorted as described above into two groups, 61 into group A and 61 into group B. Of the 122 that volunteered, three did not submit a consent form and one did not grant consent to be part of the study, as per their submitted consent form. Furthermore, 16 of those assigned to group A and 26 of those assigned to group B did not take any of the tests, and eight of group A and five of group B took only one or parts of the two tests, and of those who remained, several had corrupt recordings that were not gradable samples. In the end, there were 29 participants in group A and 25 participants in group B. Group A consisted of 14 males and 15 females. Of these, 26 of the 29 were returned missionaries, and 2 were heritage learners. Group B consisted of 10 males and 15 females. Of these, 23 of the 25 were returned missionaries, and three were heritage learners. Two of the 29 in group A and one of the 25 in group B were neither heritage learners nor returned missionaries from Japan. However, the two participants in group A had either a four-month exchange student experience or a three-month internship experience in Japan. The one participant in group B did not have any experience in Japan, but had lived in the Japan house (a foreign language housing unit provided by BYU) for two semesters. Also noteworthy is that one of the heritage learners from group A and two of the heritage learners from group B were also returned missionaries. Therefore, these heritage learners are accounted for in both the heritage learner column and also the returned missionary column though, as mentioned above, these were put in the heritage learners' stratum and not the returned missionaries' stratum. The reason for this is that the sheer volume and time of exposure to the language for heritage learners is often more significant than

an at most two year experience in the country. For a 20-year-old student, the difference in exposure to the language could be as much as ten-fold.

Table 1

Participants

	n	Male	Female	Returned Missionary	Heritage Learner
Group A	29	14	15	26	2
Group B	25	10	15	23	3

n = the number of participants in that group

Returned Missionary = the number of participants who have lived in Japan for 18 to 24 months as a missionary for the Church of Jesus Christ of Latter-day Saints

Heritage Learner = the number of participants who had and grew up with one or more native Japanese speaker as a parent

Timetable

To decrease the possible maturation effect on the study's validity, participants were required to take each of the tests in a limited time frame of one week (see Table 2), for a total of two weeks' duration for the testing. The process from recruitment to testing and data collection took place during winter semester 2015. Week 1 was the week of January 5 to January 10, week 2 was the week of January 12 to January 17, and week 3 was the week of January 19 to January 24. The reading comprehension questions and surveys from the tests were then collected in the following week, the week of January 26. The reading recordings were coded and given to two third-party graders for evaluation. The reading comprehension answer sheets were graded by the researcher. After the grading was complete and recorded, the data were analyzed by a statistics professor from BYU in April, 2015.

Table 2

Timetable

	Group A	Group B
Week 1	Recruiting Background questionnaire	Recruiting Background questionnaire
Week 2	Test 1	Test 1' (=Test 2)
Week 3	Test 2	Test 2' (=Test 1)

Instruments

The four reading scripts used in the two tests were taken from the *Japanese-Language Proficiency Test Official Practice Workbook N3* (The Japan Foundation & Japan Educational Exchanges and Services, 2014), hereafter referred to as the JLPT workbook. The Japanese Language Proficiency Test, or JLPT, consists of five levels: N1, N2, N3, N4, and N5. The easiest level is N5, and the most difficult is N1. The official website (Japanese Language Proficiency Test levels summary) describes the levels thus: “N4 and N5 measure the level of understanding of basic Japanese mainly learned in class. N1 and N2 measure the level of understanding of Japanese used in a broad range of scenes in actual everyday life. N3 is a bridging level between N1/N2 and N4/N5.” Since the study was based on the assumption that the participants were intermediate-level Japanese learners, the reading scripts from N3 were chosen for the tests, with one modification. The texts were replicated from N3’s PDF to a word document, but the furigana (phonetic kana-readings added above the kanji or Chinese characters to aid reading) were removed entirely. This is because the participants’ ability to read Chinese characters was part of the dependent variable to be measured.

There were five reading comprehension questions for each reading script (reading script 1-4). One of the five for each were taken from the JLPT workbook and used with their

corresponding script. The other four were created by the researcher in a manner similar to the question taken from the JLPT workbook. Each question was in multiple-choice format with three distractors and one correct choice. The questions and choices were all in English (see Appendix E).

Each treatment video was scripted first. A PowerPoint presentation was created accordingly to fit the script. Then, using a software program called Camtasia Studio 8, the treatment video was created while using the prepared PowerPoint presentation and script. The script was read by the researcher. In order to make each testing session take less than one hour for the participants (see Table 3), each single-treatment (ARR only or morphological and syntactic explanation only) video was made to be about two minutes in length while each combination-treatment (ARR-explanation and explanation-ARR) video was made to be about five minutes in length.

For the assisted repeated reading parts of the treatment videos, the participants heard and repeated after the audio in the video. The audio read the Japanese text and paused periodically to allow time for the participants to repeat after the recording. Due to time constraints, this process was repeated only once per repetition segment. The repetition segment was more than five words and less than ten words, so that the participants would not grow tired of having to repeat too short of a segment but would not have to strain themselves and be unable to repeat because the segment was too long. These segments were determined at the discretion of the researcher and were usually selected where the sentence ended or where there was a natural break in the sentence.

For the morphological and syntactic explanation parts of the treatment videos, the more difficult words and syntactic structures were chosen in each reading script and underlined in the

PowerPoint slides. The audio then would bring the viewers' attention to these parts and explain what those underlined segments meant in English, without providing the reading for any of the text. It was important that the reading was not provided during this treatment in order to distinguish it more clearly from the ARR treatment. After each underlined part was explained, the audio then provided a rough translation of the sentence. This was to compensate for the parts that were not highlighted previously, to emphasize further the parts that were underlined and explained, and to even out the length of the video segment with its ARR counterpart. Unless the ARR and explanation videos were the same length, one testing session would be shorter than the other and group A and group B would not end in the same amount of time. The use of translation and its effects may have been a cause for some concern and will be discussed in Chapter 4.

Procedure

The participants were asked to take two, one-hour long testing sessions in the multimedia testing lab (hereafter referred to as the testing lab) on BYU campus in Provo, Utah. With cooperation from the lab manager, the researcher was able to have the tests in the lab for the two weeks' duration of the study.

Each session consisted of two reading scripts. Each reading script had a pre-treatment recording and reading comprehension (RC) questions and a post-treatment recording and RC questions (see Table 3). At the end of the testing session, the participants were given a survey that asked them what their impressions were of the treatments. In one session, each participant submitted four sets of answers to the RC questions and four sets of recordings. Overall, each participant submitted a total of eight sets of answers to the RC questions and eight sets of recordings.

The testing was presented and the recordings were made through a program called LearningWeb. With the cooperation of the program's creator, the tests were uploaded into the program. The participants were able to access LearningWeb on their respective computers in the testing lab. The reading scripts, RC questions, and recording pages were revealed in order, during the testing session, using LearningWeb. The subjects' recordings were retrieved by the researcher after the two-week testing period had passed.

Table 3

Test Sessions

	Group A	Group B
Week 1	Recruiting Background questionnaire	Recruiting Background questionnaire
Week 2	Pre-test for reading script 1	Pre-test for reading script 1
Test	Treatment 1	Treatment 1'
Session #1	Morphological/syntactic explanation	Assisted RR method reading practice
	Post-test for reading script 1	Post-test for reading script 1
	Pre-test for reading script 2	Pre-test for reading script 2
	Treatment 2	Treatment 2'
	Assisted RR method reading practice	Morphological/syntactic explanation
	Morphological/syntactic explanation	Assisted RR method reading practice
	Post-test for reading script 2	Post-test for reading script 2
	Post-survey	Post-survey
Week 3	Pre-test for reading script 3	Pre-test for reading script 3
Test	Treatment 3	Treatment 3'
Session #2	Assisted RR method reading practice	Morphological/syntactic explanation
	Post-test for reading script 3	Post-test for reading script 3
	Pre-test for reading script 4	Pre-test for reading script 4
	Treatment 4	Treatment 4'
	Morphological/syntactic explanation	Assisted RR method reading practice
	Assisted RR method reading practice	Morphological/syntactic explanation
	Post-test for reading script 4	Post-test for reading script 4
	Post-survey	Post-survey

The four reading scripts 1-4 were given to both groups in numerical order. However, the two groups received the four treatments in a different order. Group A received morphological and

syntactic explanation as treatment 1 with reading script 1, and ARR followed by morphological and syntactic explanation as treatment 2 with reading script 2. Group B received ARR as treatment 1 with reading script 1, and morphological and syntactic explanation followed by ARR as treatment 2 with reading script 2. In the second testing week, the two groups received the set of treatments the other group received the previous week. Group A received ARR as treatment 3 with reading script 3, and morphological and syntactic explanation followed by ARR as treatment 4 with reading script 4. Group B received morphological and syntactic explanation as treatment 3 with reading script 3, and ARR followed by morphological and syntactic explanation as treatment 4 with reading script 4.

To minimize the burden on the participants, the testing sessions were limited to one hour sessions. The time allotment was estimated from entering the testing lab to exiting it (see Table 4), later referred to as time constraints.

Table 4

One Hour Estimation of Testing Sessions

Elapsed Time	Tasks Performed	Duration of Time for Task
0:00	Check in at testing lab Receive answer sheet (with the after-session survey on the reverse side) and instructions card Receive computer assignment Log in to computer and program	5 minutes
0:05	Pre-test recording for first reading script	2 minutes (max)
0:07	Computer processing/turning page time	1 minute
0:08	Pre-test RC questions for first reading script	5 minutes
0:13	Computer processing/turning page time	1 minute
0:14	First treatment video	2 minutes
0:16	Computer processing/turning page time	1 minute
0:17	Post-test recording for first reading script	2 minutes (max)
0:19	Computer processing/turning page time	1 minute
0:20	Post-test RC questions for first reading script	5 minutes
0:25	Computer processing/turning page time	1 minute
0:26	Pre-test recording for second reading script	2 minutes (max)
0:28	Computer processing/turning page time	1 minute
0:29	Pre-test RC questions for second reading script	5 minutes
0:34	Computer processing/turning page time	1 minute
0:35	Second treatment video	5 minutes
0:40	Computer processing/turning page time	1 minute
0:41	Post-test recording for second reading script	2 minutes (max)
0:43	Computer processing/turning page time	1 minute
0:44	Post-test RC questions for second reading script	5 minutes
0:49	Computer processing/turning page time	1 minute
0:50	Post-session survey	5 minutes
0:55	Log out of program and computer Submit answer sheet and return instructions card to proctor Check out of testing lab	5 minutes
1:00		

In order to have the participants perform all the tasks for this study, they had to go through a number of tasks in a short while. Most of the time allotment was estimated by experience (from the researcher taking the tests several times). The treatment videos were already made, so their durations were set. Because of the nature of the testing program used, it takes the user a click and some waiting time to start the recording and the same process to stop

the recording. Since each new task needed to open a new page, it would need some lag time. For the RC questions and post-session survey, the participants would have to refer to their answer sheet and would require some time also. To account for the processing and other preparation time, one minute was assigned as computer processing/turning page time between each page (or task).

There were five RC questions (see appendix E) and six post-session survey questions (see appendix B). Because the questions were fairly dense, both in number of words and what processing was required to answer them, these were both estimated to take about five minutes, or approximately one minute per question. The post-session survey was given less time, compared to the RC questions, since it would not require the relatively greater processing time of comprehending a foreign language script to answer the questions (though the RC questions were in English, the script from which the answers were drawn were in Japanese).

The recording time was the only time that could have been flexible, but there needed to be time set aside for the participants to prepare to start taking the test and to leave. Taking into account the other time duration estimated for the other activities and dividing the remaining eight minutes by four, the time for the recording was limited to two minutes. This time limitation to finish reading the scripts was deemed acceptable because, theoretically, setting this limit should not influence the factors that will be observed in the data, namely the reading fluency speed (in characters per second, of how many characters were read divided by how long it took to read them) and accuracy (in percentage of how many characters were read correctly, divided by how many were read). This reasoning is also discussed in the Grading section.

Grading

There were two parts that required grading in this study: the reading comprehension answer sheets and the reading recordings. The reading comprehension answers were graded by the researcher. This was deemed to be acceptable since marking right or wrong multiple-choice answers can be done objectively. To minimize the instrument decay from threatening internal validity, the pile of answer sheets to be graded were divided into four sections and graded over three days. The four sections were group A test 1, group B test 1, group A test 2, and group B test 2, and were graded in that order.

The grading of the reading recordings was given to two third-party graders. Grader A is a native Japanese speaker, a college graduate, and was able and willing to commit significant hours to grading. However, Grader A was available only for a month and was unable to finish grading all the recordings by herself. Therefore, a second grader, Grader B, was chosen. He was chosen for being a native Japanese speaker, a college graduate, and for his availability and willingness to help. The two graders were initially given 440 recordings to grade. However, several recordings were inaudible or missing due to technical issues. This problem, unfortunately, was not revealed until the graders had finished grading. Because, as previously mentioned, some volunteers did not have a full set of recordable data, there were 54 participants at the end of the grading stage. Each of the 54 participants had recorded eight, gradable reading recordings. Therefore, of the 440 that were graded, 432 were incorporated into the study.

Due to Grader A's time restrictions, she was not able to do the grading for all 440 recordings by herself. Because there were two graders, there needed to be a method to check the inter-rater reliability between them. In order to do this, after all the recordings had been assigned to one of the two graders, five of Grader B's assigned recordings were randomly chosen and also

assigned to Grader A. Likewise, five of Grader A's assigned recordings were randomly chosen and also assigned to Grader B. In all, Grader A was given 350 (345 of her own exclusive recordings to grade, and five recordings that were duplicates of those assigned to Grader B), and Grader B was given 100 (95 of his own exclusive recordings to grade, and five recordings duplicated from Grader A's). The graders were not notified of this measure.

The two graders were trained in how to grade the recordings in a one-hour training session. At the time of the training, the graders were given two laminated sheets of the four reading scripts, a CD of the recordings they were to grade, some dry-erase markers, and erasers. The instructions given to the graders were as follows:

1. Listen to the recording and mark all the characters misread or skipped on the laminated sheet. Count and note the number of characters misread or skipped. If multiple errors of the same word occur, write which words and how many were missed.
2. Note how long the reading took (two minutes were the maximum set on the software used). If the reader was able to complete the reading, put down the time that the recording ended. If the reader was unable to complete the reading, put down how many characters they did read within the allotted time.
3. Record #1 and #2 in a Google sheet in the corresponding rows (see Appendix F).

It was determined that the grading for reading fluency would be based on accuracy and speed, taking into account the methods used by Shen and Jiang (2013) and Jeon (2012). Though Shen and Jiang (2013) also measured the prosody in segmentation, this seemed less applicable in Japanese than Chinese that was the medium of their study, and Jeon (2012) cited that prosody was largely subjective. Therefore, the graders were asked to evaluate only the accuracy and to

note the speed. The maximum length of the recording was set to be two minutes through the testing software. The reason for the two minutes time limit is because the participants had a set amount of tasks to accomplish within the one hour testing session, and two minutes was the most time that could be allotted to each recording within the one hour block. Two minutes were deemed to be enough for the nature of this study because of the factors, noted above, that would be graded. Assuming that the speed of reading is largely consistent for a participant, the number of words read in two minutes divided by two minutes should be similar to the number of words in the script divided by how long it took to finish reading the script.

The graders were to note the time if the participant was able to read all the reading script before the two minutes elapsed and to note the number of characters that were read in this two minute period if they did not finish reading all of the script. The laminated scripts given to the graders noted the number of characters in increments. With these noted times or noted number of characters, through computations in a spreadsheet, the researcher was able to divide the number of characters read (either the total number of characters in the script or the number noted by the grader) by the time it took to read (120 seconds or the time noted by the grader). In this manner, each recording's speed, in characters per second, was computed and recorded.

Both graders finished their allotted grading within the deadline given, which was a month, and reported back to the researcher. Though the one-month deadline was given by the researcher, it was also the Grader A's availability deadline. Because Grader A would not be available after one month, Grader B was given the same deadline. Each recording took an average of 5-6 minutes to grade. The inter-rater reliability for the two graders will be discussed in the following chapter.

Chapter 4: Results and Discussions

Results and Data Analysis

The reading fluency (RF), as mentioned above, was graded on reading accuracy (per character) and reading speed. The reading comprehension (RC) score was given as a percentage of how many correct answers were chosen out of the five pre- and post- test scores. To avoid subconscious bias, the reading fluency was graded by two third-party graders. On the other hand, since the reading comprehension questions were simple to grade (right or wrong), these were graded by the researcher. The descriptive statistics can be seen in Tables 5 and 6.

Table 5

Descriptive statistics of group A

	Script#1 (205)			Script#1			Script#2 (160)			Script#2		
	Pre-test			Post-test			Pre-test			Post-test		
	RF acc.	RF spd.	RC	RF acc.	RF spd.	RC	RF acc.	RF spd.	RC	RF acc.	RF spd.	RC
Mean	82.20 4%	2.564	69.65 5%	88.62 9%	2.921	77.24 1%	80.05 8%	1.703	43.44 8%	89.80 6%	2.303	69.65 5%
Median	80.00 0%	2.412	80.00 0%	87.80 5%	2.929	80.00 0%	78.75 0%	1.649	40.00 0%	89.37 5%	2.192	60.00 0%
Mode	79.02 4%	2.531	80.00 0%	99.02 4%	3.534	100.0 00%	78.75 0%	2.133	40.00 0%	100.0 00%	1.839	60.00 0%
Low	72.19 5%	1.708	20.00 0%	77.07 3%	1.708	0.000 %	66.25 0%	0.975	0.000 %	70.62 5%	1.441	60.00 0%
High	99.02 4%	3.942	100.0 00%	100.0 00%	4.184	100.0 00%	100.0 00%	2.857	100.0 00%	100.0 00%	3.478	100.0 00%
Range	26.82 9%	2.234	80.00 0%	22.92 7%	2.475	100.0 00%	33.75 0%	1.882	100.0 00%	29.37 5%	2.037	40.00 0%
Variance	0.509 %	0.451	28.44 2%	0.487 %	0.445	34.10 2%	0.797 %	0.223	34.57 8%	0.568 %	0.302	7.753 %
St. Dev.	7.261 %	0.683	24.27 3%	7.099 %	0.679	26.57 8%	9.086 %	0.480	26.76 3%	7.671 %	0.559	12.67 2%

	Script#3 (213)			Script#3			Script#4 (195)			Script#4		
	Pre-test			Post-test			Pre-test			Post-test		
	RF acc.	RF spd.	RC	RF acc.	RF spd.	RC	RF acc.	RF spd.	RC	RF acc.	RF spd.	RC
Mean	84.55 4%	2.202	62.06 9%	88.79 5%	2.591	80.00 0%	92.87 4%	2.464	66.89 7%	97.41 8%	3.204	76.55 2%
Median	84.03 8%	2.130	60.00 0%	87.79 3%	2.566	80.00 0%	92.30 8%	2.349	60.00 0%	97.94 9%	3.305	80.00 0%
Mode	81.69 0%	2.130	80.00 0%	87.79 3%	2.566	100.0 00%	88.20 5%	2.120	60.00 0%	100.0 00%	2.378	80.00 0%
Low	63.36 6%	0.842	0.000 %	76.05 6%	1.467	20.00 0%	83.07 7%	1.639	0.000 %	92.30 8%	2.191	40.00 0%
High	98.59 2%	4.019	100.0 00%	99.06 1%	4.019	100.0 00%	100.0 00%	4.149	100.0 00%	100.0 00%	4.239	100.0 00%
Range	35.22 5%	3.177	100.0 00%	23.00 5%	2.552	80.00 0%	16.92 3%	2.510	100.0 00%	7.692 %	2.048	60.00 0%
Variance	0.690 %	0.433	41.85 5%	0.372 %	0.390	28.96 6%	0.213 %	0.405	32.10 5%	0.055 %	0.390	20.78 5%
St. Dev.	8.451 %	0.670	29.44 5%	6.211 %	0.635	24.49 5%	4.697 %	0.648	25.78 8%	2.392 %	0.636	20.75 0%

The RF accuracy is given in percentage of how many characters were read correctly divided by how many were attempted

The RF speed is given by # of characters read divided by how many were read

The RC score is given in percentage of how many questions were answered accurately divided by 5 (the number of questions)

The length of each script is given in number of characters after each script

Table 6

Descriptive statistics of group B

	Script#1 (205)			Script#1			Script#2 (160)			Script#2		
	Pre-test			Post-test			Pre-test			Post-test		
	RF acc.	RF spd.	RC	RF acc.	RF spd.	RC	RF acc.	RF spd.	RC	RF acc.	RF spd.	RC
Mean	84.43 7%	2.569	80.00 0%	86.79 0%	3.282	91.20 0%	82.76 6%	2.221	51.20 0%	91.17 4%	2.638	73.60 0%
Median	83.41 5%	2.595	80.00 0%	86.82 9%	3.417	100.0 00%	81.58 3%	2.065	60.00 0%	95.62 5%	2.424	80.00 0%
Mode	77.56 1%	2.847	80.00 0%	96.09 8%	3.475	100.0 00%	75.62 5%	1.333	60.00 0%	98.75 0%	2.192	80.00 0%
Low	74.63 4%	1.058	40.00 0%	74.14 6%	1.400	40.00 0%	56.25 0%	1.283	0.000 %	72.48 3%	1.242	40.00 0%
High	99.51 2%	4.881	100.0 00%	99.51 2%	5.256	100.0 00%	98.12 5%	4.103	100.0 00%	100.0 00%	4.211	100.0 00%
Range	24.87 8%	3.823	60.00 0%	25.36 6%	3.856	60.00 0%	41.87 5%	2.819	100.0 00%	27.51 7%	2.969	60.00 0%
Variance	0.460 %	0.533	17.60 0%	0.518 %	0.771	11.32 8%	0.925 %	0.595	35.32 8%	0.768 %	0.526	10.75 2%
St. Dev.	6.922 %	0.745	19.14 9%	7.343 %	0.896	15.36 2%	9.825 %	0.788	27.12 9%	8.942 %	0.740	14.96 7%

	Script#3 (213)			Script#3			Script#4 (195)			Script#4		
	Pre-test			Post-test			Pre-test			Post-test		
	RF acc.	RF spd.	RC	RF acc.	RF spd.	RC	RF acc.	RF spd.	RC	RF acc.	RF spd.	RC
Mean	87.25 5%	2.441	59.20 0%	92.19 0%	3.060	84.00 0%	94.12 2%	2.953	64.00 0%	97.72 0%	3.559	79.20 0%
Median	88.26 3%	2.290	60.00 0%	95.30 5%	3.132	80.00 0%	94.87 2%	2.708	60.00 0%	98.97 4%	3.545	80.00 0%
Mode	79.81 2%	2.663	80.00 0%	89.20 2%	2.803	80.00 0%	94.35 9%	2.708	80.00 0%	99.48 7%	3.482	80.00 0%
Low	68.06 7%	0.992	0.000 %	65.38 5%	0.867	40.00 0%	85.64 1%	1.358	40.00 0%	89.23 1%	1.358	40.00 0%
High	98.59 2%	4.019	100.0 00%	100.0 00%	5.325	100.0 00%	99.48 7%	5.132	100.0 00%	100.0 00%	5.132	100.0 00%
Range	30.52 4%	3.027	100.0 00%	34.61 5%	4.458	60.00 0%	13.84 6%	3.773	60.00 0%	10.76 9%	3.773	60.00 0%
Variance	0.460 %	0.520	48.76 8%	0.646 %	0.942	12.80 0%	0.147 %	0.852	16.00 0%	0.074 %	0.792	21.56 8%
St. Dev.	6.922 %	0.736	31.87 5%	8.203 %	0.991	16.33 0%	3.912 %	0.942	18.25 7%	2.779 %	0.908	21.19 7%

The RF accuracy is given in percentage of how many characters were read correctly divided by how many were attempted

The RF speed is given by # of characters read divided by how many were read

The RC score is given in percentage of how many questions were answered accurately divided by 5 (the number of questions)

The length of each script is given in number of characters after each script

In Tables 5 and 6, the reading fluency (RF) accuracy is given as a percentage of how many characters were read correctly, divided by how many were attempted. The reading fluency (RF) speed is given by the number of characters read (whether correctly or incorrectly), divided by the time it took to read, in seconds. The reading comprehension (RC) scores are given as a percentage of how many questions were answered correctly, divided by five (the total number of RC questions). Notice that there is a noticeable increase in all three scores from each pre-test to post-test scores.

The results were analyzed using ANOVA where the independent variables were the treatments and their interactions with each other and the dependent variables were the reading fluency (accuracy and speed) and reading comprehension. The analyses of the results are shown in Tables 7, 8, and 9.

Table 7

Analysis of RF accuracy

Treatment	Estimate	Standard Error	DF	t value	p
1 ARR	0.03336	0.006641	153	5.02	<.0001*
2 Explanation	0.05636	0.006641	153	8.49	<.0001*
3 ARR-Exp	0.06684	0.006641	153	10.06	<.0001*
4 Exp-ARR	0.06423	0.006641	153	9.67	<.0001*

ARR=Assisted Repeated Reading, Exp=Explanation=Morphological and Syntactic explanation

*=statistical significance

Table 8

Analysis of RF speed

Treatment	Estimate	Standard Error	DF	t value	p
1 ARR	0.03336	0.006641	153	5.02	<.0001*
2 Explanation	0.05636	0.006641	153	8.49	<.0001*
3 ARR-Exp	0.06684	0.006641	153	10.06	<.0001*
4 Exp-ARR	0.06423	0.006641	153	9.67	<.0001*

ARR=Assisted Repeated Reading, Exp=Explanation=Morphological and Syntactic explanation

*=statistical significance

Table 9

Analysis of Reading Comprehension

	Treatment	Estimate	Standard Error	DF	t value	p
1	ARR	0.6983	0.1844	153	3.79	0.0002*
2	Explanation	0.8147	0.1844	153	4.42	<.0001*
3	ARR-Exp	1.0302	0.1844	153	5.59	<.0001*
4	Exp-ARR	0.8664	0.1844	153	4.70	<.0001*

ARR=Assisted Repeated Reading, Exp=Explanation=Morphological and Syntactic explanation

*=statistical significance

As indicated in the tables, each treatment showed a significant gain from the participant's pre-test scores to their post-test scores, but there was not a statistically significant difference between the two treatments or combination of the treatments (discussed later in this chapter). This result answers the first three research questions.

Research question #1 asked, "How effective is Assisted Repeated Reading (ARR), where the learners repeatedly practice reading the same passage in the target language while having access to a native speaker's audio example, in improving reading fluency for intermediate Japanese learners?" Since there was a significant gain in all categories for all methods, including ARR, it can be said that ARR is effective in improving reading fluency for intermediate Japanese learners. However, because there was not a significant difference between the methods, it cannot be conclusively stated that ARR is the sole reason for the improvement in reading fluency. It is possible that the mere repeated exposure (second time seeing the same text in the post-test, as opposed to it being their first time in the pre-test, aside from seeing the text on screen for 2-5 minutes during the treatment video) could have been responsible for the statistically significant improvement that was observed. However, the participants had equal or greater exposure to the text during their treatment in regards to repeated exposure to the text. Taking this into account and including the exposure to the text during the treatment, this factor of repeated exposure that

was not part of the treatment can have at most an equal but likely a lesser effect on the result as the treatment.

A greater concern may be the fact that the participants had the opportunity to repeat after the recording only once during the ARR treatment as opposed to the three or more repetitions that usually compose an ARR treatment. Therefore, this treatment may not have had the full effect of this method. This was due to time and technological constraints. It was not possible to make more opportunities to repeat without having this single treatment (ARR) video become twice or three times longer than its counterpart video of morphological and syntactic explanation. Due to time constraints, the single treatment videos needed to be about two minutes in length. And due to the researcher's lack of technological expertise, a method to contain more repetitions within the same length of time could not be devised. Because there was only one repetition instead of several, this treatment was not as effective as it could have been. This could be one of the reasons why there was not a significant difference shown between the treatments, specifically why ARR did not have a more significant influence on the participants' reading fluency gain than the other treatments.

Research question #2 asked, "How effective is morphological and syntactic explanation in improving reading comprehension for intermediate Japanese learners?" Since there was a significant gain in all categories for all methods, including morphological and syntactic explanation, it can be said that morphological and syntactic explanation does seem to be an effective method to improve reading comprehension. However, similar to the ARR above, because there was not a significant difference between the methods, it cannot be conclusively stated that morphological and syntactic explanation is the sole reason for the improvement in reading comprehension. It is possible that the mere repeated exposure (second time seeing the

same text and questions in the post-test, as opposed to it being their first time in the pre-test, though the RC questions were in different order between the pre- and post-tests) could have been responsible for the statistically significant improvement that was observed. It is also important to keep in mind that the text was on the screen for the duration of the treatment video (2-5 minutes depending on the treatment). So, similar to the discussion above, if the exposure to text during treatment was part of the treatment, the other exposure to text could be considered an equal or lesser factor in the result.

Another factor to be considered is the fact that the treatment included a translation of each sentence, along with the morphological and syntactic explanation. This translation could have played a larger role in helping the participants' comprehension of the text than the explanation could have by itself. However, because the translation was with the explanation in the same treatment, it cannot be determined how much effect it might have had on the results. There were too few perfect (100%) scores on the RC post-tests to say that the translation clearly had too much influence, but the effect of the translation may have been lessened because of the order of the testing, i.e. the participants may have forgotten the translation by the time they finished their RF test and were to take the RC test.

Research question #3 asked, "Will Assisted Repeated Reading also influence reading comprehension and will morphological and syntactic explanation also improve reading fluency?" As previously stated, reading fluency was measured in two parts, through accuracy in reading characters and through speed. Since both the reading comprehension and fluency scores improved from pre- to post- test scores, there does seem to be some crossover benefit. ARR not only helped reading fluency but also seemed to help reading comprehension in intermediate level Japanese learners, and morphological and syntactic explanation not only helped reading

comprehension but also seemed to help reading fluency. Though this, again, cannot be conclusively stated to be a fact because there was not a statistical significance between the different methods, it can be a good indication to note.

In research questions 1-3, the fact that the methods did not show significant differences in their results has some underlying implications that will be seen in the analysis of the data. This is an unexpected result since it can imply that the participants may have been at least partially improving just by repeated exposure to the text and not necessarily due to the treatments employed. However, it could also imply that both methods were equally helpful in improving the participants' reading fluency and reading comprehension. There is not a way to determine the reason for this result, but several factors may have contributed to the outcome. It may be due to the fact that the pre-test scores were already very high. The pre-test scores averaged 86.02% in RF accuracy, 2.40 characters per second in RF speed, and 3.10 out of 5, or 62.0% in RC scores. The fact that there was not too much room to improve may have been a strong factor in this result. Another factor is that there were only five RC questions, so each question was worth 20%. This was unavoidable due to time constraints (see Chapter 3) and length of the script, but the RC scores were strongly influenced by this setting. Though these high pre-test scores can be assumed to be a higher contributing factor, it is possible that if there were more RC questions, at least the RC scores would not have had as many unexpected post-test scores (i.e. negative gains from pre- to post-test RC scores).

Since there were two graders, inter-rater reliability was taken into account. As stated above in the grading section of Chapter 3, grader A was given 350 recordings and grader B was given 100 recordings to grade. Five of grader A's assigned recordings were also given to grader B (included in his 100) and five of grader B's assigned recordings were also given to grader A

(included in her 350) to check for inter-rater reliability. The five recordings from each assigned recordings were chosen randomly by the researcher.

Inter-rater reliability can be measured by agreement (Ato, Lopez, & Benavente, 2011) or correlation (Kayapinar, 2014). Ato, Lopez, & Benavente (2011) chose to analyze the agreement coefficients such as Bennet's sigma, Scott's pi, Cohen's kappa, and Gwet's gamma, in their study. However, since the data for inter-rater reliability for this study only has ten recordings and two graders (Table 7), analyzing this data based on pure agreement (whether grader A and grader B's corresponding ratings were a match) was not feasible. However, because there were 345 that were assigned to grader A and only 95 to grader B, there were not enough data to make an agreement analysis feasible, unless all or most of 95 that were assigned to grader B were also given to grader A. Keeping in mind that it took an average of five minutes to grade one recording, even an increase of six would equal 30 more minutes of the graders' time. The chosen graders were not able to handle that many more assigned recordings (see Grader section in Chapter 3).

As shown in Table 10, it is clear that the graders were not in complete agreement with each other in score or time, but they were only one or two points and/or seconds apart from each other for most of the disagreements. Therefore, it was decided that the inter-rater reliability for this study would be measured in terms of correlation.

Table 10

Inter-rater Reliability data

Recordings	Grader A's score (# of characters misread or skipped)	Grader B's score (# of characters misread or skipped)	Grader A's time in seconds	Grader B's time in seconds
A1	9	10	78	79
A2	13	15	84	84
A3	8	10	65	65
A4	12	14	80	80
A5	22	23	68	69
B1	21	23	100	101
B2	24	26	51	51
B3	22	28	87	88
B4	29	32	94	95
B5	1	1	60	61

A1-5=Recordings originally assigned to grader A; B1-5=Recordings originally assigned to grader B

To judge the inter-rater reliability between the two graders, the Pearson r correlation was computed for both sets of data, the number of characters misread or skipped and the time in seconds. The result of this computation can be seen in Table 11. The correlations between the two graders were very high. The correlation between graders for the number of characters misread or skipped had an r value of 0.9904, and the time in seconds had an r value of 0.9995.

Table 11

Inter-rater Reliability Correlation

	# of characters misread or skipped	Time in seconds
Pearson r	0.9904	0.9995

From this analysis of correlation for inter-rater reliability, it can be stated that though the two graders did not have high agreement, they were able to produce ratings with similar values that showed high correlation to each other. The use of two graders, if not carefully controlled for subjectivity, could be “highly vulnerable to observer differences” (Fraenkel, Wallen, & Hyun, 2012, p. 159) and could influence the internal validity of the study. However, in this case, the two

graders were able to produce compatible results and can be said to have strong inter-rater reliability.

Though touched upon earlier in the chapter while discussing the other three research questions, the answer to research question four required further analysis. Research question four stated: when the methods are employed in a different order, will there be a difference in the students' level of improvement in their reading fluency and reading comprehension?

Similar to the first three questions, the data (Tables 12-14) were not able to show a statistical significance between the two different treatments (whether ARR came first or second in the combination treatments). However, it was noted that when the two methods were employed together (regardless of order), the participants generally had higher gains than when single methods were employed alone (Table 15). Though not statistically significant or found to be a consistent pattern, this result has some implications that will be discussed in the next section.

Table 12

Differences between Treatments in Analysis of RF accuracy

	Treatment	Estimate	Standard Error	DF	t value	p
1	ARR-exp to ARR	0.03348	0.008199	153	4.08	<.0001*
2	Exp-ARR to ARR	0.03087	0.008199	153	3.77	0.0002*
3	Exp to ARR	0.02300	0.008199	153	2.81	0.0057*
4	Exp-ARR to ARR-exp	-0.00261	0.008199	153	-0.32	0.7506
5	Exp to ARR-exp	-0.01048	0.008199	153	-1.28	0.2033
6	Exp to Exp-ARR	-0.00787	0.008199	153	-0.96	0.3389

ARR=Repeated Reading, Exp=Explanation=Morphological and Syntactic explanation

*=statistical significance

Table 13

Differences between Treatments in Analysis of RF speed

	Treatment	Estimate	Standard Error	DF	t value	p
1	ARR-exp to ARR	0.05243	0.08464	153	0.62	0.5366
2	Exp-ARR to ARR	0.04462	0.08464	153	0.53	0.5989
3	Exp to ARR	-0.06673	0.08464	153	-0.79	0.4317
4	Exp-ARR to ARR-exp	-0.00781	0.08464	153	-0.09	0.9266
5	Exp to ARR-exp	-0.1192	0.08464	153	-1.41	0.1612
6	Exp to Exp-ARR	-0.1113	0.08464	153	-1.32	0.1903

ARR=Repeated Reading, Exp=Explanation=Morphological and Syntactic explanation

*=statistical significance

Table 14

Differences between Treatments in Analysis of Reading Comprehension

	Treatment	Estimate	Standard Error	DF	t value	p
1	ARR-exp to ARR	0.3319	0.2501	153	1.33	0.1865
2	Exp-ARR to ARR	0.1681	0.2501	153	0.67	0.5025
3	Exp to ARR	0.1164	0.2501	153	0.47	0.6424
4	Exp-ARR to ARR-exp	-0.1638	0.2501	153	-0.65	0.5135
5	Exp to ARR-exp	-0.2155	0.2501	153	-0.86	0.3902
6	Exp to Exp-ARR	-0.05172	0.2501	153	-0.21	0.8364

ARR=Repeated Reading, Exp=Explanation=Morphological and Syntactic explanation

*=statistical significance

Table 15 shows the gains that were observed in each treatment. For more details of the treatments and their order, see Table 3 in Chapter 3. For more details of the descriptive statistics, from which the figures in Table 15 were taken, see Tables 5 and 6. Observe that the trend is not consistent, but in general, the combination treatment yields higher gains.

Table 15

Gains Observed in Each Treatment

	Group A	Group B
Exp. alone	6.426% RF acc. (Script 1) 0.356 RF speed 7.586% RC	4.935% RF acc. (Script 3) 0.619 RF speed 24.800% RC
ARR-Exp.	9.749% RF acc. (Script 2) 0.600 RF speed 26.207% RC	3.597% RF acc. (Script 4) 0.606 RF speed 15.200% RC
ARR alone	4.241% RF acc. (Script 3) 0.389 RF speed 17.931% RC	2.353% RF acc. (Script 1) 0.713 RF speed 11.200% RC
Exp.-ARR	4.545% RF acc. (Script 4) 0.740 RF speed 9.655% RC	8.408% RF acc. (Script 2) 0.417 RF speed 22.400% RC

Exp. = morphological and syntactic explanation, RF = reading fluency, acc. = accuracy

ARR= assisted repeated reading, RC = reading comprehension

Since all the methods produced statistically significant gains in reading fluency and comprehension, it can be said that all of the methods were effective in helping the learners' reading fluency and comprehension. Though not statistically significant, the morphological and syntactic explanation treatment had generally higher gains than its counterpart of ARR. Similarly, the ARR-Exp (ARR before morphological and syntactic explanation) treatment had generally higher gains than its counterpart of Exp-ARR (morphological and syntactic explanation before ARR).

Chapter 5: Summary and Conclusion

Summary

There was a total of 54 participants in this study that took place during winter semester 2015 at Brigham Young University, Provo. The participants were divided in two groups—29 in group A and 25 in group B. These participants participated in two one-hour testing sessions in a computer lab, and their pre-test reading fluency (measured by both accuracy and speed) and reading comprehension scores were compared with the corresponding post-test scores. The results showed statistically significant improvement in all the treatments employed. Also, even though it did not show a statistically significant difference between the treatments, there were some noticeable differences seen between the various treatments.

This study was not able to isolate the separate effects of the methods to show conclusively that assisted repeated reading (ARR) is an effective teaching method for increased reading fluency, or that morphological and syntactic explanation is an effective teaching method for increased reading comprehension. Likewise, it could not show conclusively that ARR can increase reading comprehension, nor that morphological and syntactic explanation can increase reading fluency. Upon the point of having a combination of the two methods, the study was also not able to show conclusively which order is more effective than the other. However, this could also imply that both methods, in either order, were somewhat effective in increasing the learners' reading fluency and reading comprehension. In fact, the combination treatments (of ARR and morphological and syntactic treatment together) yielded a statistically significant difference in gains for reading fluency accuracy than ARR alone. However, this may be more due to how these methods were implemented in the treatments (i.e. the ARR had only one repetition but the morphological and syntactic explanation also included English translation of the text, as noted in

Chapter 4) and may not be a reflection on the methods themselves. Future studies should take this into account and adjust their instruments accordingly.

Though the results were not statistically significant and therefore not conclusive, there are several good implications that can be derived from this study. There were indications that ARR, morphological and syntactic explanation, and their two combinations were effective in causing gains in reading fluency and reading comprehension. Though there were not statistically significant differences in the four methods, the results did show that:

1. All the methods showed significant gains in both reading fluency and reading comprehension.
2. The combination of the two methods generally caused larger gains in both reading fluency and reading comprehension than the single treatments.
3. Morphological and syntactic explanation treatment alone had generally higher gains than its counterpart of ARR alone.
4. The ARR-Exp (ARR before morphological and syntactic explanation) treatment had generally higher gains than its counterpart of Exp-ARR (morphological and syntactic explanation before ARR).

Implications

There is much more that can be done in this field of research. To improve foreign language education and specifically to improve learners' reading fluency and reading comprehension, more research on this topic are required for more conclusive results.

Though the results did not show a significant difference in favor of any one treatment, they did show significant improvement with all the treatments. The following paragraphs outline some of the important implications that can be derived from the results of this study.

The intermediate-level Japanese learners that participated in this study were all observed to have made significant gains through the treatments given to them. This implies that learners were able to make significant gains in reading fluency and comprehension by receiving either a single treatment of about two minutes or a combination of treatments of about five minutes. The pedagogical implication, then, is that by applying as little as two minutes of class time or homework time (with guidance) to the learners' reading fluency and/or comprehension, learners can make significant gains in reading fluency and comprehension. From this study, it is not clear whether the gains are only short-term or could extend to long-term gains, but further studies may be able to add to these results. However, it can be said that it may be advisable for language instructors, especially Japanese teachers, since that was the language of focus in this study, to add to their classroom routine, a few minutes of reading and explaining practice of the same text. The same text can be used more than twice, for this practice. However, again, the gains could be short-term or long-term. A longitudinal study on this topic may add to these results.

Some of the pedagogical implications that can be derived from these results are: (a) that learners can benefit with gains in reading fluency and reading comprehension from even minimal (2-5 minutes') practice or explanation, and (b) that applying both methods yields generally higher gains in reading fluency and reading comprehension than one method alone. With these results, teachers of Japanese and other foreign languages may be able to consider using both methods every class to improve their students' reading fluency and reading comprehension.

Limitations of the Study

Due to the time limitations imposed and in order to not put too much burden on the participants, the tests were set up to last approximately one hour for each testing session. In order to collect all the necessary data in this one-hour block, there were several tasks to be performed

within this time frame (Table 4). Also, because of this time constraint, there could not be too many reading comprehension (RC) questions. The RC questions required reading and thinking that could not be time-limited by the software. Therefore, the RC questions were limited to just five, which made one missed question equal to a loss of 20% on the participants' RC scores. This could be a strong factor in why the RC scores fluctuated so much. There were several students that got 20% fewer points in their post-test RC score, than their corresponding pre-test RC, because they got one more question incorrect in the post-test. This may indicate that these students were guessing and had not actually comprehended the text. This is another factor that may have influenced the RC scores.

Though there were some discrepancies in the length (the shortest being script 2 with 160 characters and the longest being script 3 with 213 characters), all the scripts were taken from JLPT N3 workbook and thereby, should be approximately the same level of difficulty. However, the nature of the questions, especially the four that were created by the researcher, may be another contributing factor to the RC fluctuations.

The time constraint also contributed to why there was only one repetition per reading script instead of multiple, as is more common when applying ARR method (Taguchi, Gorsuch, Takayasu-Maass, & Snipp, 2012). This was also likely a large factor in why there was not a distinguishable gain difference between the methods. If there were multiple repetitions, there likely would have been higher gains on reading fluency scores in the post-test scores when the ARR method was employed. These higher gains would have resulted in a larger difference in gains from the morphological and syntactic explanation method.

There was also a time constraint on when the testing lab could be used. Because the testing lab is in high demand from other departments, this study had to take place before the

midterms began, in the first month of the semester. This seemingly contributed to fewer volunteers and subsequently fewer participants because there was less time to publicize and to gather volunteers, and the students may not have been as motivated by extra credit at the beginning of the semester and therefore likely did not follow through with their agreement to take the tests (see participants' section).

Suggestions for Future Research

Though it may be more stressful on the participants, it would be better to challenge them with higher level tasks so there will be wider range of differences in pre- to post- test scores. It will also be advisable to have more participants that are more traditional students who have learned their Japanese mostly in academic settings rather than through experience abroad and/or at home. ACTFL guidelines (Liskin-Gasparro J.E., 1982, as cited in <http://www.languagetesting.com/how-long-does-it-take>) estimate that 480 hours of language instruction or training are required, on average, to achieve an intermediate-low level of proficiency in Japanese. If the learner has had extensive exposure to the language or learning opportunity outside of the classroom, their time duration in language instruction become more complicated to estimate. Therefore, it will be more difficult to estimate their level of proficiency without actually testing their proficiency level with a proficiency assessment. In future studies, it may be useful to have some proficiency assessments in the beginning of the study.

Furthermore, though this was not implemented in this study due to time and technological constraints, the ARR method usually calls for more than one repetition per reading segment. If replicating this study, it is recommended that the participants be given more opportunities to repeat, e.g. three times or more as Taguchi, Gorsuch, Takayasu-Maass, & Snipp (2012) chose to do in their study.

Research of L2 language learners' reading fluency and reading comprehension, especially their relationship with each other, is a topic that has incited recent interest in the language teaching field. This may be because producing more fluent readers who can comprehend well also is a goal that many language teachers can and do aspire toward. It is an interesting and worthy topic that requires further research.

References

- Ato, M., Lopez, J. J., & Benavente, A. (2011). A Simulation Study of Rater Agreement Measures with 2x2 Contingency Tables. *Psicologica: International Journal of Methodology and Experimental Psychology*, 32, (2), 385-402.
- Fox, B. J. (2006). Fluency contributes to comprehension. Retrieved from HYPERLINK "http://www.education.com/print/fluency-contributes-comprehension/"
<http://www.education.com/print/fluency-contributes-comprehension/>.
- Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education* (8th edition). New York: McGraw-Hill.
- Hu, M., & Nation, I. S. P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a foreign language*, 13, (1), 403-430.
- Japanese Language Proficiency Test levels summary. (2015). Retrieved from <https://www.jlpt.jp/e/about/levelsummary.html>.
- Jeon, E.H. (2012). Oral reading fluency in second language reading. *Reading in a foreign language*, 24, (2), 186–208.
- Jodai, H., & Tahriri, A. (2011). Reading rate and comprehension. *Modern journal of language teaching methods*, 1, (3), 122-131.
- Jorden, E.H., & Walton, A.R. (1987). Truly foreign languages: Instructional challenges. *The annals of the American academy of political and social science*, 490, (1), 110-124.
- Kayapinar, U. (2014). Measuring Essay Assessment: Intra-rater and Inter-rater Reliability. *Eurasian Journal of Educational Research*, 57, 113-136.

- Kim, Y-S., Wagner, R.K., & Lopez, D. (2012). Developmental relations between reading fluency and reading comprehension: A longitudinal study from grade 1 to grade 2. *Journal of experimental child psychology*, 113, 93-111.
- LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive psychology*, 6, (2), 293-323.
- Language Testing International. (2014). How long does it take to become proficient? Retrieved from HYPERLINK "<http://www.languagetesting.com/how-long-does-it-take>" <http://www.languagetesting.com/how-long-does-it-take> .
- Lightbown, P.M., & Spada, N. (2013). *How languages are learned* (4th edition). Oxford: Oxford University Press.
- Liskin-Gasparro, J. E. (1982). *ETS Oral Proficiency Testing Manual*. Princeton, N.J.: Educational Testing Service. Abstract retrieved from <http://www.languagetesting.com/how-long-does-it-take>
- Murray, M. S., Munger, K. A., & Clonan, S. M. (2012). Assessment as a strategy to increase oral reading fluency. *Intervention in school and clinic*, 47, (3), 144-151.
- Nation, I.S.P., & Newton, J. (2009). *Teaching ESL/EFL listening and speaking*. New York, NY: Routledge.
- Ning, C. (2009). Engaging a “Truly Foreign” language and culture: China through Chinese film. *Educational Perspectives*, 42, 29-35.
- Ocana, H. V. (2011). Effects of extended explicit systematic phonics instruction on adult L2 reading fluency. (Masters’ thesis). Retrieved from Electronic Theses & Dissertations in BYU Harold B. Lee digital collections. (etd 4498).

- Rossiter, M. J., Derwing, T. M., Manimtim, L. G., & Thomson, R. I. (2010). Oral fluency: The neglected component in the communicative language classroom. *The Canadian modern language review*, 66, 4, 583-606.
- Sakade, F. (2003). *A guide to reading & writing in Japanese* (3rd edition). North Clarendon, VT: Tuttle Publishing.
- Schmidt, R.W. (1992). Psychological mechanisms underlying second language fluency. *Studies in second language acquisition*, 14, (04), 357-385.
- Shen, H. H., & Jiang, X. (2013). Character reading fluency, word segmentation accuracy, and reading comprehension in L2 Chinese. *Reading in a Foreign Language*, 25, (1), 1-25.
- 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, (4), 30-55.
- The Japan Foundation, & Japan Educational Exchanges and Services. (2014). *Japanese-Language Proficiency Test Official Practice Workbook N3* (3rd Edition). Tokyo, Japan: Bonjinsha.
- Tong, X., Deacon, S. H., & Cain, K. (2013). Morphological and syntactic awareness in poor comprehenders: another piece of the puzzle. *Journal of learning disabilities*, 47, (1), 22-33.

Appendices

Appendix A

Background questionnaire

Name: _____

1. Which Japanese class(es) are you enrolled in this semester?

2. What Japanese classes have you taken previously and how well did you do in those classes, on average?

Classes taken: _____

Average grade (circle one): -----A B C D

E

3. Have you had extensive exposure to Japanese language, outside of what was required for the above listed classes (e.g. experience living in Japan, living with a Japanese relative, attending Japanese school, having a Japanese exchange student stay with your family)?

Activity

Time duration

a. _____

b. _____

c. _____

If more, continue on reverse side of this questionnaire.

4. Do you have family members (in the immediate family that you have grown up in) that you spoke/used Japanese with, on a regular basis?

Yes/No -----If yes, who (by relationship)? -----

Appendix B

Post-session survey

Date _____

Circle one for each treatment today:

Treatment 1: reading practice by repeating / explanation of the reading passage

Treatment 2: reading practice—explanation / explanation—reading practice

1. Did either task present any difficulty for you? If so, which task? What made it difficult?

Yes/No (Circle one and explain below if yes)

2. How did you feel you did on the pretests?

3. How did you feel you did on the posttests?

4. Do you feel there was a perceivable difference in how well you did on the pretest/posttest? Why?

Yes/No (Circle one and explain reason below if yes)

5. Which do you think contributed to better reading fluency for you: today's treatment 1 or treatment 2?

Treatment 1/ Treatment 2 (Circle one or explain below)

6. Which do you think contributed to better reading comprehension for you: treatment 1 or treatment 2?

Treatment 1/ Treatment 2 (Circle one or explain below)

Appendix C

Instructions for Participants

1. Sit down at an open computer with A) an answer sheet, and B) this instructions card.	2. Write the date, your name and group on the Answer sheet (if you're not sure about group, look on the backside of this instructions card).
3. Locate and open the Learning web	4. Choose and click on "NetRecorder"
5. Type in your net ID (you won't receive credit if I don't know who participated).	6. When it says your name, click "yes."
7. Choose "Traci Okawa thesis for Intermediate Japanese readers"	8. Click on the test you are taking (if you don't remember the group, look on the back. Week 1—January 12-17 Week 2—January 19-24
	9. Follow the instructions on screen. Voice test, etc. (You will need to push record for a second & stop, to proceed with the test, even on non-recording tasks.) close video screens
	10. Exit the program, log out of computer, and give both sheets to attendant on your way out. Thank you for your cooperation. Traci (Okawa) Andreason

Appendix D

Reading scripts

Reading #1

The following is a memorandum left on Ms. Nakamura's desk by the professor :

中村さん、

おはようございます。

昨日プリンターが故障したので、川名電気に修理を頼みました。修理の人は、10時に来てくれるそうです。わたしは授業があるので、修理の人が来たら、プリンターの場所に案内をお願いします。

故障の内容ですが、印刷するときに紙にインクの汚れがついてしまいます。実際に印刷した紙がプリンターのところに置いてありますから、それを修理の人に見せて説明してください。

修理がすんだら、午後のゼミの資料を人数分用意しておいてください。

下田

Reading #2

The following is an e-mail sent to all of the students attending Dr. Kawashima's seminar:

川島先生のお別れ会について、詳しいことが決まりましたので、お知らせします。7月8日（金）までに参加するかどうかを返信してください。

日時：8月10日（水） 午後7時～9時

会場：レストラン「春」

会費：3,000円

記念品代：500円（記念品としてネクタイを贈りたいと思います）

会費と記念品代は会場で集めます。

参加できない人は、記念品代だけを7月中に払ってください。

大田

Reading #3

先週、うれしいことがあった。

支店で難しい問題が発生し、広島に出張することになった。三日目にやっと解決でき、ほっとしてホテルに戻ったのだが、荷物を整理したとき、間違えて重要な書類を捨ててしまった。しかし、気がつかずにその日夜遅く東京に帰ってきた。翌朝気がついて、あわててホテルに電話をしたら、すぐに書類を見つけてくれた。そのホテルでは、客がチェックアウトしたあとも部屋のゴミはもう一泊させるのだそうだ。客のことをよく考えたサービスだと感心し、本当にうれしかった。

Reading #4

子どものころのことを思い出してみてください。雲が動物の形に見えたり、壁のしみや汚れが顔に見えたりしたことはありませんか。また、雨や風の音を聞いて、音楽のようだと感じたことがある人もいるかもしれません。では、大人になった今はどうでしょう。多くの方が、大人になると、そのように感じる「子どもの心」をなくしてしまいます。「子どもの心」を持ち続け、それによって感じたものを音楽や絵で表すことのできる人が芸術家なのではないでしょうか。

Appendix E

Reading Comprehension Questions

Reading 1 Pre-test:

- 1) What can you reasonably assume Ms. Nakamura's relationship is to the professor?
 - a. His student
 - b. His mentor teacher
 - c. His secretary
 - d. His co-worker
- 2) According to the text, what should Ms. Nakamura do when the repairperson comes?
 - a. Go to the classroom to get Dr. Shimoda.
 - b. Put the printed paper by the printer.
 - c. Explain how the printer broke to the repairperson.
 - d. Guide the repair person to the printer.
- 3) Where is the printed paper that Ms. Nakamura needs to show the repairperson?
 - a. On her desk
 - b. In the professor's classroom
 - c. By her computer
 - d. By the printer
- 4) Why can't the professor meet the repairperson himself?
 - a. He is teaching a class
 - b. He is preparing for a class
 - c. He is meeting with students
 - d. He is taking a day off
- 5) What is wrong with the printer?
 - a. It smears the printed letters
 - b. It puts ink spots on the paper
 - c. The paper jams
 - d. The print is unclear

Reading 1 Post-test:

- 1) What is wrong with the printer?
 - a. The paper jams
 - b. The print is unclear
 - c. It puts ink spots on the paper
 - d. It smears the printed letters
- 2) Why can't the professor meet the repairperson himself?
 - a. He is taking a day off
 - b. He is teaching a class
 - c. He is meeting with students
 - d. He is preparing for a class
- 3) According to the text, what should Ms. Nakamura do when the repairperson comes?
 - a. Explain how the printer broke to the repairperson.
 - b. Put the printed paper by the printer.
 - c. Guide the repairperson to the printer.
 - d. Go to the classroom to get Dr. Shimoda.
- 4) Where is the printed paper that Ms. Nakamura needs to show the repair person?
 - a. In the professor's classroom
 - b. By the printer
 - c. On her desk
 - d. By her computer
- 5) What can you reasonably assume Ms. Nakamura's relationship is to the professor?
 - a. His secretary
 - b. His co-worker
 - c. His mentor teacher
 - d. His student

Reading 2 Pre-test:

- 1) Who can you reasonably assume the sender of this e-mail to be?
 - a. A secretary
 - b. Dr. Kawashima
 - c. A student
 - d. A colleague of Dr. Kawashima
- 2) What is this e-mail about?
 - a. Dr. Kawashima's birthday party
 - b. Dr. Kawashima's goodbye party
 - c. Dr. Kawashima's welcome party
 - d. A celebratory party hosted by Dr. Kawashima
- 3) What do attendees need to bring to this event?
 - a. 500 yen and a tie
 - b. A gift and 3,000 yen
 - c. 3,500 yen
 - d. 3,000 yen, plus cost for food
- 4) What are the instructions for those who cannot attend?
 - a. Pay 3,500 yen by the time of the party
 - b. Pay 500 yen within the month of July
 - c. Pay 3,000 yen within the month of July
 - d. RSVP by the middle of July that they can't attend
- 5) From the content, we can assume that the topic in the e-mail (Dr. Kawashima's party) has . . .
 - a. Never been discussed before
 - b. Been discussed and particulars decided upon
 - c. Been discussed before but not had particulars decided upon
 - d. Been discussed only by Dr. Kawashima and family

Reading 2 Post-test:

- 1) What is this e-mail about?
 - a. Dr. Kawashima's welcome party
 - b. A celebratory party hosted by Dr. Kawashima
 - c. Dr. Kawashima's goodbye party
 - d. Dr. Kawashima's birthday party
- 2) What do attendees need to bring to this event?
 - a. 3,500 yen
 - b. 3,000 yen, plus cost for food
 - c. A gift and 3,000 yen
 - d. 500 yen and a tie
- 3) Who can you reasonably assume the sender of this e-mail to be?
 - a. A colleague of Dr. Kawashima
 - b. A student
 - c. a secretary
 - d. Dr. Kawashima
- 4) From the content, we can assume that the topic in the e-mail (Dr. Kawashima's party) has . . .
 - a. Been discussed before but not had particulars decided upon
 - b. Been discussed only by Dr. Kawashima and family
 - c. Never been discussed before
 - d. Been discussed and particulars decided upon
- 5) What are the instructions for those who cannot attend?
 - a. Pay 3,000 yen within the month of July
 - b. Pay 500 yen within the month of July
 - c. RSVP by the middle of July that they can't attend
 - d. Pay 3,500 yen by the time of the party

Reading 3 Pre-test:

- 1) Where does the author usually live?
 - a. Tokyo
 - b. Higashikyo
 - c. Hiroshima
 - d. Koshima
- 2) Where did he go on his business trip?
 - a. Tokyo
 - b. Higashikyo
 - c. Hiroshima
 - d. Koshima
- 3) What was the problem he discovered upon his return?
 - a. He had lost some important documents
 - b. He threw away some important documents
 - c. He forgot some important documents
 - d. Some important documents were stolen
- 4) What was the hotel's service that impressed him?
 - a. Their housekeeping organized his documents for him
 - b. They remembered him by name and room number
 - c. They let guests stay one additional night at a discount
 - d. They keep the trash in the room after the guest checks out
- 5) What was the "happy thing" mentioned in the first sentence?
 - a. He was able to resolve the difficult problem at work
 - b. The hotel organized his documents for him
 - c. The hotel hadn't discarded his documents
 - d. He was able to stay in the same room for one more night

Reading 3 Post-test:

- 1) What was the hotel's service that impressed him?
 - a. They remembered him by name and room number
 - b. They let guests stay one additional night at a discount
 - c. They keep the trash in the room after the guest checks out
 - d. Their housekeeping organized his documents for him
- 2) What was the problem he discovered upon his return?
 - a. Some important documents were stolen
 - b. He forgot some important documents
 - c. He had lost some important documents
 - d. He threw away some important documents
- 3) What was the "happy thing" mentioned in the first sentence?
 - a. The hotel hadn't discarded his documents
 - b. He was able to stay in the same room for one more night
 - c. He was able to resolve the difficult problem at work
 - d. The hotel organized his documents for him
- 4) Where did he go on his business trip?
 - a. Hiroshima
 - b. Higashikyo
 - c. Koshima
 - d. Tokyo
- 5) Where does the author usually live?
 - a. Hiroshima
 - b. Higashikyo
 - c. Koshima
 - d. Tokyo

Reading 4 Pre-test:

- 1) Who is the audience that this author has in mind?
 - a. Children
 - b. Adults
 - c. Teenagers
 - d. Retirees
- 2) What does the author seem to believe most people have experienced during their childhood?
 - a. Saw faces in smudges and smears on the walls
 - b. Smudged and smeared the walls with their faces
 - c. Heard music in the winds and the trees
 - d. Found animal shapes in the smudges on the walls
- 3) What does the author seem to believe that some but not too many people would have experienced during their childhood?
 - a. Smudged and smeared the walls with their faces
 - b. Heard music in the winds and the trees
 - c. Found animal shapes in the clouds
 - d. Saw faces in smudges and smears on the walls
- 4) According to the passage, this author believes that the “heart of a child” is . . .
 - a. Often found in children
 - b. Seldom found in adults
 - c. Often lost in childhood
 - d. Found only in artists
- 5) According to the author, artists are people who can use the “heart of a child” to . . .
 - a. Relate to nature
 - b. Receive artistic inspirations
 - c. Express their feelings
 - d. Keep childlike imaginations

Reading 4 Post-test:

- 1) According to the author, artists are people who can use the “heart of a child” to . . .
 - a. Express their feelings
 - b. Receive artistic inspirations
 - c. Keep childlike imaginations
 - d. Relate to nature
- 2) What does the author seem to believe that some but not too many people would have experienced during their childhood?
 - a. Heard music in the winds and the trees
 - b. Found animal shapes in the clouds
 - c. Saw faces in smudges and smears on the walls
 - d. Smudged and smeared the walls with their faces
- 3) According to the passage, this author believes that the “heart of a child” is . . .
 - a. Found only in artists
 - b. Often lost in childhood
 - c. Seldom found in adults
 - d. Often found in children
- 4) What does the author seem to believe most people have experienced during their childhood?
 - a. Heard music in the winds and the trees
 - b. Saw faces in smudges and smears on the walls
 - c. Found animal shapes in the smudges on the walls
 - d. Smudged and smeared the walls with their faces
- 5) Who is the audience that this author has in mind?
 - a. Teenagers
 - b. Children
 - c. Retirees
 - d. Adults

Appendix F

Reading Fluency Grading Criteria

Track #	Characters misread	Time/# of words within 2 min	Notes
Example	40	1:17	故障(2)、修理(5)、印刷(2)
Practice 1			
Practice 2			

Appendix G

Answer Sheet

Name: _____

Date: _____ Group: A / B

Answer sheet

1) Treatment 1 pre-test (Please don't change answer after treatment)

Fill in the answer bubble that best matches the question on screen.

1. a b c d
2. a b c d
3. a b c d
4. a b c d
5. a b c d

3) Treatment 2 pre-test (Please don't change answer after treatment)

Fill in the answer bubble that best matches the question on screen.

1. a b c d
2. a b c d
3. a b c d
4. a b c d
5. a b c d

2) Treatment 1 post-test

Fill in the answer bubble that best matches the question on screen.

1. a b c d
2. a b c d
3. a b c d
4. a b c d
5. a b c d

4) Treatment 2 post-test

Fill in the answer bubble that best matches the question on screen.

1. a b c d
2. a b c d
3. a b c d
4. a b c d
5. a b c d