The Role of Inference in Second Language Reading Comprehension: Developing Inferencing Skill Through Extensive Reading

Sayako Niwa

Follow this and additional works at: https://scholarworks.umass.edu/masters_theses_2

Recommended Citation
https://scholarworks.umass.edu/masters_theses_2/792

This Open Access Thesis is brought to you for free and open access by the Dissertations and Theses at ScholarWorks@UMass Amherst. It has been accepted for inclusion in Masters Theses by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.
THE ROLE OF INFERENCE
IN SECOND LANGUAGE READING COMPREHENSION:
DEVELOPING INFERENCING SKILLS THROUGH EXTENSIVE READING

A Thesis Presented

by

SAYAKO NIWA

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

MASTER OF ARTS

May 2019

Department of Languages, Literatures, and Cultures

East Asian Languages and Cultures

Japanese
THE ROLE OF INFERENCE
IN SECOND LANGUAGE READING COMPREHENSION:
DEVELOPING INFERENCING SKILLS THROUGH EXTENSIVE READING

A Thesis Presented

by

SAYAKO NIWA

Approved as to style and content by:

___________________________________________
Yuki Yoshimura, Chair

___________________________________________
Bruce Baird, Member

___________________________________________
David K. Schneider, Unit Director
East Asian Languages and Cultures Program
Department of Languages, Literatures and Cultures

___________________________________________
Robert G. Sullivan, Chair
Department of Languages, Literatures and Culture
ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to my advisor, Professor Yuki Yoshimura for her continuous support. I cannot imagine completing this study without her. Her guidance helped throughout the research and writing of this thesis. Thank you very much for always being a supportive advisor and mentor.

I would also like to thank Professor Bruce Baird for his insightful comments and support for completing this study.
ABSTRACT
THE ROLE OF INFERENCE IN SECOND LANGUAGE READING COMPREHENSION:
DEVELOPING INFERENCING SKILLS THROUGH EXTENSIVE READING
MAY 2019
SAYAKO NIWA, B.A., NANZAN UNIVERSITY
M.A., UNIVERSITY OF MASSACHUSETTS AMHERST
Directed by: Professor Yuki Yoshimura
The purpose of this study is to determine whether extensive reading has positive effects on developing inferencing skills. Extensive reading is a language learning method of reading large amounts of comprehensible texts. This method limits the use of dictionaries while reading; therefore, extensive readers have greater practice in dealing with unfamiliar words than non-extensive readers. One of the ways to deal with unfamiliar words is to infer the meaning of the word using contextual clues. Knowing how to infer the meaning of unknown words is a helpful skill for language learners. Due to the fact that extensive readers have a greater practice in dealing with unknown words, this study examines whether there are any differences in the precision of inferencing skills between extensive readers and non-extensive readers. There were 39 participants analyzed in this study, 28 non-extensive readers and 11 extensive readers. The results showed that extensive reading has positive effects on language learners’ inferencing skills. In terms of accuracy, we could not see a statistical difference; however, the extensive readers had a higher percentage in accurately inferring the word meaning. In terms of the use of knowledge sources, extensive readers were able to choose the appropriate knowledge
source when inferring the target word. These results indicate that extensive reading can enhance language learners’ inferencing skills.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>ix</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2. THEORETICAL FRAMEWORK</td>
<td>3</td>
</tr>
<tr>
<td>2.1. Lower Level Processes</td>
<td>3</td>
</tr>
<tr>
<td>2.1.1. Word Recognition</td>
<td>3</td>
</tr>
<tr>
<td>2.1.2. Syntactic Parsing</td>
<td>7</td>
</tr>
<tr>
<td>2.2. Higher Level Processes</td>
<td>9</td>
</tr>
<tr>
<td>2.3. L1 and L2 Differences in Reading</td>
<td>11</td>
</tr>
<tr>
<td>2.4. Inference</td>
<td>13</td>
</tr>
<tr>
<td>3. LITERATURE REVIEW</td>
<td>20</td>
</tr>
<tr>
<td>3.1. Prior Research on Inference in Reading</td>
<td>20</td>
</tr>
<tr>
<td>3.1.1. A study on ESL Learners in Canada</td>
<td>20</td>
</tr>
<tr>
<td>3.1.2. A Study on ESL Learners in Japan</td>
<td>21</td>
</tr>
<tr>
<td>3.1.3. A Study on Japanese Language Learners</td>
<td>23</td>
</tr>
<tr>
<td>3.2. Extensive Reading</td>
<td>24</td>
</tr>
<tr>
<td>3.3. Significance of This Study</td>
<td>27</td>
</tr>
<tr>
<td>3.4. Research Questions</td>
<td>28</td>
</tr>
</tbody>
</table>
4. METHODS AND PARTICIPANTS ................................................................. 30
   4.1. About the Participants of This Research ....................................... 30
   4.2. Research Design ........................................................................ 30
   4.3. Data Collection ......................................................................... 34
   4.4. Data Analysis ........................................................................... 36

5. ANALYSIS AND RESULTS ................................................................. 39
   5.1. Analyzing Research Question 1 .................................................... 39
   5.2. Interpretation of the Analysis (RQ1) ............................................. 40
   5.3. Analyzing Research Question 2 .................................................... 42
       5.3.1. The Usage of Knowledge Sources to Infer Target Words in the 1st Reading Material (Yūdachi) .................................................... 44
       5.3.2. The Usage of Knowledge Sources to Infer Target Words in the 2nd Reading Material (Amedama) ............................................ 46
       5.3.3. The Amount of Knowledge Sources Utilized ......................... 47
   5.4. Interpretation of the Analysis (RQ2) .............................................. 48

6. CONCLUSION ..................................................................................... 51
   6.1. Limits of This Study ................................................................... 51
   6.2. Conclusion .................................................................................. 52

APPENDICES .......................................................................................... 55

   APPENDIX A: SURVEY SECTION OF THE RESEARCH ....................... 56
   APPENDIX B: READING SECTION OF THE RESEARCH ...................... 60
   APPENDIX C: IRB .............................................................................. 70

REFERENCES ......................................................................................... 72
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The Nassaji Model Categorization (adapted from Nassaji, 2003)</td>
<td>16</td>
</tr>
<tr>
<td>2. The Wesche &amp; Paribakht Model Categorization (adapted from Wesche &amp; Paribakht, 2009)</td>
<td>17</td>
</tr>
<tr>
<td>3. Comparison of Intensive Reading and Extensive Reading</td>
<td>25</td>
</tr>
<tr>
<td>4. Morphemes used in the First Reading Material “Yūdachi”</td>
<td>32</td>
</tr>
<tr>
<td>5. Morphemes used in the Second Reading Material “Amedama”</td>
<td>32</td>
</tr>
<tr>
<td>6. List of Pseudowords and How It was Determined (1st reading material)</td>
<td>33</td>
</tr>
<tr>
<td>7. List of Pseudowords and How It was Determined (2nd reading material)</td>
<td>33</td>
</tr>
<tr>
<td>8. 3-point Scale Used to Determine Inferencing Accuracy</td>
<td>37</td>
</tr>
<tr>
<td>9. Taxonomy of Five Knowledge Sources</td>
<td>38</td>
</tr>
<tr>
<td>10. Inferencing Accuracy for the First Reading Material</td>
<td>39</td>
</tr>
<tr>
<td>11. Inferencing Accuracy for the Second Reading Material</td>
<td>39</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Processes in Word Recognition (adapted from Adams, 1990)</td>
<td>6</td>
</tr>
<tr>
<td>2. General Parsing Flow when Reading Japanese Sentences</td>
<td>9</td>
</tr>
<tr>
<td>3. Non-Extensive Readers- Percentages of Knowledge Sources Used</td>
<td>45</td>
</tr>
<tr>
<td>to Infer Target Words (1st Reading Material)</td>
<td></td>
</tr>
<tr>
<td>4. Extensive Readers- Percentages of Knowledge Sources Used to</td>
<td>45</td>
</tr>
<tr>
<td>Infer Target Words (1st Reading Material)</td>
<td></td>
</tr>
<tr>
<td>5. Non-Extensive Readers- Percentages of Knowledge Sources Used</td>
<td>46</td>
</tr>
<tr>
<td>to Infer Target Words (2nd Reading Material)</td>
<td></td>
</tr>
<tr>
<td>6. Extensive Readers- Percentages of Knowledge Sources Used to</td>
<td>47</td>
</tr>
<tr>
<td>Infer Target Words (2nd Reading Material)</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 1
INTRODUCTION

When you encounter an unknown word while reading a text, what do you do? There are many ways to overcome this problem: one way is to infer the meaning of the word using contextual cues and your own knowledge. “Lexical Inferencing” is an important ability in reading comprehension; however, it is difficult to practice in second language reading. Although we go through this process without any effort when reading in our native language, why does it become difficult in second language reading? There are many factors which make inferring in a second language text difficult. The main reason is, largely, because most learners have not had enough practice in reading second language texts.

Japanese language text books for students do not include many reading materials. This is because they are primarily focused on developing communication skills. The main structure of a textbook is introducing a model conversation and explaining the grammar and vocabulary used in the conversation. Japanese language learners normally learn through this process; however, when they become advanced language learners, they are suddenly required to read complex reading materials. It is very difficult because they do not have enough reading experience, and as such, the learners will have difficulty in reading fluently, and tend to rely on dictionaries. Similarly, developing inferencing skills is also difficult.

However, Extensive Reading has started to intervene in some language learning environments. Extensive reading is a language learning method that requires learners to
read as much as they can without using a dictionary. Even elementary level language learners can learn to read fluently through this method. When we go back to the point why inferring in second language reading is difficult, it may be because of the lack of reading experience. Extensive reading is a method which learners read a lot and can start from any language levels. Therefore, they will have more reading experience than just going through a language textbook. Furthermore, because extensive reading limits the use of dictionaries, learners have to deal with words that they do not know on their own. Considering these points, we can assume that extensive reading may be one way to develop inferencing skills. This study will examine the relationship between extensive reading and inferencing skills.
CHAPTER 2
THEORETICAL FRAMEWORK

2.1 Lower Level Processes

2.1.1 Word Recognition

Lower level processes include letter, word, and sentence level comprehension. Word recognition is the first process which readers execute when reading a text and is a process which readers extract lexical information from written words on the texts. (Koda, 2005) It is widely accepted that learners who have better ability in reading comprehension can also recognize words rapidly and automatically; in short, fluent readers are fluent word recognizers. There are researchers who differentiate between the terms word recognition and decoding, for example, Crandall et al. (n.d.) defines word recognition as a process to see and recognize a word immediately without effort, and decoding as an act to use the alphabet and its letter-sound to produce the pronunciation of the word. However, Koda (2005) mentions that word recognition and decoding are often used interchangeably. Therefore, this study adopts Koda’s stance and use the term “word recognition” consistently. Word recognition is contained with various lower level processes such as orthographic processing, phonological processing, and semantic processing. (reviewed by Grabe, 2009) Each of the processes will be described in the next paragraphs.

When reading a text, readers first see the letters/words on the text to gain information. This process, to recognize letters and words in the text, is called
orthographic processing. Orthographic processing is the first process which readers use when reading a text and is the only process which receives information directly from the text. (Adams, 1990) In the case of reading Japanese texts, orthographic processing is responsible for recognizing the three orthographies used in Japanese; kanji, hiragana, and katakana.

After recognizing letters/words through orthographic processing, readers generate the sound of its letters/words. This is called phonological processing; the process which is “responsible for mapping the letters into spoken equivalent.” (Adams, 1990) Phonological processing does receive outside information; however, phonological processing only receives speech information. Although it does not accept information from reading texts, when readers try to comprehend a text, it is often the case to read the letters/words with its sound in their mind. Readers generate the sound(s) of the letters/words they recognized in their mind because it allows them to access their mental lexicon quickly. (Frost, 1998)

Semantic processing uses the information gained through orthographic processing and phonological processing. Through orthographic and phonological processes, the readers understood the sequences of letters and its pronunciation. Using this information, the readers try to understand what the word means. This process, to figure out the meaning of the word, is semantic processing. (Adams, 1990) To sum up the three processes, for example, when readers encounter a word “走る” in Japanese texts, the readers first recognize it as a word that combines kanji and hiragana by going through orthographic processing. The readers then recognizes the word that can be pronounced “hashi-ru” by accessing their mental lexicon by going through phonological processing.
After this process, the readers again have access to their mental lexicon and understand that this word means “to run” by going through semantic processing.

These three processes are essential in word recognition and reading comprehension; however, understanding the context is necessary to comprehend a text. Adams (1990) described the process to construct “an on-going understanding of the text” as a context processor. When readers read the sentence “あめがふる” in a Japanese text, they first look at the letters and figure out that the pronunciation of this sentence is “Ame ga furu.” If the readers only go through until semantic processing, the readers may be confused because Ame has two different meanings in Japanese; rain and candy drop. If the readers use context processor, they can solve the semantic ambiguity because the verb “furu” means “to fall” for meteorological things. Thus, the readers can verify that the noun Ame in this context should mean “rain.” As we can see in the example, the context processor is responsible for selecting the appropriate meaning in the context and creating coherence to the text.

Adams (1990) illustrated the four processes below in Figure 1.
As we can see in Figure 1, all the processes work together as a team. It keeps on receiving and returning information back and forth to comprehend a text. It is also important to understand that all these processes need to cooperate with their mental lexicon in order to accurately comprehend texts. As it is described when explaining how to comprehend a sentence “Ame ga furu,” the reader needs to use their own knowledge to figure out what the used letters are, how the letters can be pronounced, what each word means, and what meaning fits the context. When reading text in the first language (L1), accessing their mental lexicon is generally easier and accurate; however, for second
language (L2) reading, there are cases in which the learners cannot figure out the meaning of the sentence because their mental lexicon is not equipped enough. When this is the case, semantic processing and context processor cannot be processed while word recognition still takes place. (Adams, 1990)

The readers need to perform these word recognition processes automatically in less than a quarter of a second for each word to achieve fluent reading. When fluent readers encounter a text, they automatically go through the processes to recognize and understand the word. In order to recognize words automatically, accurate word recognition processes and well-developed lexical entries are necessary. However, automaticity is not a skill everyone owns naturally. Elementary readers first go through the word recognition processes manually. By experiencing and practicing large amounts of reading and having meaningful input, readers eventually acquire the skill to go through the processes without making efforts and become able to go through the word recognition processes automatically. Thus, developing automaticity in word recognition requires a large amount of reading. (reviewed by Grabe, 2009)

### 2.1.2 Syntactic Parsing

Syntactic parsing is also a lower level process that happens simultaneously with the word recognition process. Syntactic parsing is a process to access meaningful information in their mental lexicon by using information from the word recognition process. (reviewed by Grabe, 2009) When readers encounter a sentence “The police man protected a little girl with a teddy bear,” there are two ways to comprehend. If the teddy bear is the modifier, a little girl had the teddy bear, but if the police man was the modifier,
the police man had the teddy bear. Syntactic parsing is responsible for determining which
to modify. Thus, syntactic parsing plays a major role in determining the meaning of a
sentence. (Koda, 2005)

In Japanese, syntactic parsing plays more crucial roles because syntactic
structures are entirely different from English. Although English has spaces between each
word, Japanese do not have any spaces. This makes it challenging to understand where
each word starts and ends in a sentence. For example, “私もももが食べられない”
(watashi mo momo ga taberarenai) in Japanese, meaning “I also cannot eat a peach” in
English requires special attention. The Japanese sentence is confusing to distinguish each
word, especially if the reader does not have enough Japanese vocabulary knowledge. The
general flow of parsing to comprehend a sentence is first to divide the sentence into each
morpheme, and then combine some of the divided morphemes into semantically
meaningful chunks (Iwatate, 2012). Figure 2 uses the above sentence to show how the
readers divide the sentence into chunks in detail.
2.2 Higher Level Processes

Higher level processes are responsible for comprehending texts in a broader level than lower level processes such as comprehending discourse level, paragraph level, and even longer texts. It comprehends by integrating syntactic information and its meaning of the discourse and paragraph which the readers encounter. Further, higher level processes also perform to monitor their reading processes and metacognition.
processes. (Terauchi, 2010) Metacognition processes are self-reflection processes, which allows the readers to look at the processes objectively which they themselves are going through while reading. These processes are often automatic for fluent readers; however, the automaticity to read fluently in higher level processes is not as important as it is in lower level processes. This is because the automaticity of higher-level processes cannot be achieved without the automaticity of lower-level processes. (reviewed by Grabe, 2009)

There are two main component abilities in the higher level processes; text model of comprehension and situation model of reader interpretation. (Grabe, 2009) Text model is a process which understands texts by using the information within the text. It integrates the main concept and the subconcept of the sentences to figure out the central ideas and to understand the meaning of the sentence or longer texts. The situation model involves an entirely different reading process from the text model. Situation model is a process which understands the text by using the information readers own. The information comes from, for example, how the readers look at the text, what kind of attitude the readers have towards the theme of texts, what kind of thoughts the readers have when reading similar texts in the past, how readers evaluate the text itself, and so on. (Terauchi, 2010)

A text model is prioritized when the reader has little background knowledge of the text information as they cannot fully interpret to what the writer is trying to say nor do they have any opinions about the text. On the other hand, a situation model is preferred when the reader has strong background knowledge. Although the proportion of the uses of these two models differs by the type of texts the reader’s encounter, the two models are normally used together while the reader tries to comprehend a text. The readers use the text model to understand what the writer aimed to tell readers through the
text and keep the relevant information active in their network. However, none of the readers respond the same way as others when they read the same text. Therefore, readers use their own ideas, knowledge, interpretations of the text, which is the situation model. As we can see, both the text model and the situation model are essential for reader comprehension. (reviewed by Grabe, 2009)

Readers who read in their second language tend to rely on the situation model rather than the text model when reading challenging texts. This is because they do not have enough vocabulary to catch information only from what the text explains. The situation model requires the readers to interpret what the text provides, and therefore, is being used more so the reader can use their knowledge and ideas to understand the text. However, the use of the situation model by L2 readers does not guarantee that the reader is comprehending the text accurately because L2 readers lacks usage of the text model process. When readers lack usage of the text model process, they are interpreting without understanding the writer’s aim of the text. (reviewed by Grabe, 2009)

2.3 L1 and L2 Differences in Reading

There are other elements that influence reading comprehension in addition to the detailed reading processes explained in the previous sections. Two of the elements are L1 and L2 differences. As mentioned in the introduction, it is common to have different feelings towards reading L1 texts and L2 texts. This is because reading L2 texts is a completely different task from reading L1 texts. There are various reasons why L2 reading entails a more complex task. Among the various differences, this study points out the three differences in L1 and L2 reading as follows:
1. Processing differences
2. Knowledge differences
3. Differences in the amount of practice

The L1 and L2 readers’ differences of processing is noticeable. When comparing the amount of time it takes when reading in L1 and L2, it is usually the case that it takes more time when reading L2 texts. One of the reasons why reading L2 texts are time-consuming is because the readers have less developed lexical knowledge. Because the readers do not have much lexical knowledge, it slows down processes such as word recognition, syntactic processing, and accessing their mental lexicon. Another reason is that there are more processes to go through when reading L2 texts such as going through both L1 and L2 resources. When reading in L1, the readers go through the processes only in their L1; however, when reading in L2, readers go through the processes by using both their L1 and L2. When comparing L1 and L2’s process of reading, reading L2 texts requires an additional process because L2 readers need to go through both L1 and L2 resources, thus, it takes more time to comprehend L2 texts. (reviewed by Grabe, 2009)

Knowledge is also a different element in their L1 and L2 reading. The lexical, grammatical, and discourse knowledge which learners utilize is different between L1 and L2 readers, which is particularly so for beginners of the target language. In L1, children orally have enough lexical, grammatical, and discourse knowledge to read when they start reading. However, in the case of L2, the readers are expected to comprehend a text while they are still in the process to acquire vocabulary. In L1, the readers can start strengthening their reading ability after building up their lexical knowledge, while L2 readers have to gain both a variety of linguistic knowledge and the reading ability at the
same time. (reviewed by Grabe, 2009)

As mentioned previously, reading comprehension improves with practice, thus, it is important to read as much as they can to enhance reading comprehension. Although this is the case, when we compare L1 reading and L2 reading in terms of reading amount, the lack of reading practice in L2 leads to greater difficulties in reading comprehension. As for L1, there are a multitude of chances to practice reading because it is used in their daily life and readers tend to choose to read in L1 when there is more than one language available. On the other hand, for L2, reading practice is very limited. For language learners who learn the language outside of a country where the target language is used, it is common to only have reading practice during class time and while doing homework. This is certainly not enough to enhance their reading skills effectively. Thus, lack of practice is another reason why L2 learners have more difficulty to improve their reading skills compared to reading L1.

The three differences above — processing difference, knowledge difference, and reading amount differences in L1 and L2 reading — are some of the reasons why L2 reading is a more time-consuming and a difficult task for language learners. However, as mentioned continuously in this study, reading comprehension can be improved with practice. With extensive practice, learners can achieve smoother processing and improve their working knowledge used when reading.

2.4 Inference

As mentioned in earlier sections, there are many processes and different types of knowledge involved in reading comprehension such as understanding lexical information,
grammatical information, and so on. Inferencing is also one of the processes involved in reading comprehension. In a broader sense, readers infer during the normal course of reading to establish coherence to the text-meaning construction (Koda, 2005). However, inference is also a process that occurs when readers encounter words that they do not know in the text. While the readers’ goal is to comprehend and understand the text in its entirety, the readers also give attention to a particular word and then infer the meaning of the word in order to comprehend the text, in other words, lexical inferencing. Because L2 readers do not have enough vocabulary knowledge, L2 readers have more chances to practice inferring the meanings of words they do not know. As it is outlined in the previous sections, reading comprehension skills have a tendency to improve by practice. If this is true, lexical inferencing skills will also improve in the same way.

Haastrup (1991) defines lexical inferencing as the process that makes “informed guesses as to the meaning of a word in the light of all available linguistic cues in combination with the learner’s general knowledge of the world, her awareness of the context and her relevant linguistic knowledge.” To identify an appropriate meaning of a word, the readers need to find useful cues from the surrounding context clues, understand the flow of the text (especially for narrative texts), and draw on previous knowledge. Although the process of lexical inferencing is similar for L1 and L2 readers, it is true that the L1 readers have advantages in both accuracy and speed because they have richer linguistic and cultural knowledge for comprehending texts. Wesche and Paribakht (2009) argues that lexical inferencing ability can be a particularly valuable tool for L2 readers and can help them improve in L2 reading skills. This is because there are often times when L2 readers need to deal with unfamiliar words. If L2 readers improve their lexical
inferencing skills, they will enhance their reading fluency, which will support their academic learning.

When readers try to infer word meanings, they use various types of knowledge both inside and outside of the text. The use of readers’ knowledge sources has been a major topic of discourse within the researching of lexical inference. Previous studies have shown that readers use their language knowledge and world knowledge during the process of comprehension by gaining information from content and linguistic cues in the text. It is also known that when reading an L2 text, the readers use both their L1 knowledge and L2 knowledge, and that readers’ L1 language and educational histories affect the way readers use their knowledge source. (Wesche and Paribakht, 2009)

Through the process of analyzing studies in the past and the think-aloud protocols process, Nassaji (2003) categorized five knowledge sources which L2 readers use when inferring words that they do not understand. He defined the knowledge sources as “instances when the learner made an explicit reference to a particular source of knowledge” and the categories of knowledge sources include: grammatical knowledge, morphological knowledge, knowledge of L1, world knowledge and discourse knowledge. From this point in this study, this classification categorized by Nassaji will be called the Nassaji model.
Paribakht will be called the Wesche & Paribakht model. From this point, this taxonomy organized by Wesche and Paribakht will be called the Wesche & Paribakht model.
<table>
<thead>
<tr>
<th>Linguistic Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2-based sources</td>
</tr>
<tr>
<td>• <strong>Word Knowledge</strong></td>
</tr>
<tr>
<td>- <em>Word Association</em>: Association of the target word with another familiar word or a network of words.</td>
</tr>
<tr>
<td>- <em>Word Collocation</em>: Knowledge of words that frequently occur with the target word.</td>
</tr>
<tr>
<td>- <em>Word Morphology</em>: Morphological analysis of the target word based on knowledge of grammatical inflections, stem and affixes.</td>
</tr>
<tr>
<td>- <em>Word Form (written)</em>: Knowledge of formal (orthographic or phonetic) similarity between the target word and another word resembling it.</td>
</tr>
<tr>
<td>• <strong>Sentence Knowledge</strong></td>
</tr>
<tr>
<td>- <em>Sentence Meaning</em>: The meaning of part or all of the sentence containing the target word.</td>
</tr>
<tr>
<td>- <em>Sentence Grammar</em>: Knowledge of the syntactic properties of the target word, its speech part and word order constraints.</td>
</tr>
<tr>
<td>- <em>Punctuation</em>: Knowledge of rules of punctuation and their significance.</td>
</tr>
<tr>
<td>• <strong>Discourse Knowledge</strong></td>
</tr>
<tr>
<td>- <em>Discourse Meaning</em>: The perceived general meaning of the text and sentences surrounding the target word (i.e. beyond the immediate sentence that contains the target word).</td>
</tr>
<tr>
<td>- <em>Formal Schemata</em>: Knowledge of the macro structure of the text, text types and discourse patterns and organization.</td>
</tr>
<tr>
<td>- <em>Text Style &amp; Register</em>: Knowledge of stylistic and register variations in word choice.</td>
</tr>
<tr>
<td>L1-based sources</td>
</tr>
<tr>
<td>• <strong>L1 Collocation</strong></td>
</tr>
<tr>
<td>Knowledge of words in L1 that have collocational relationship with the L1 equivalent of the target word, assuming that the same relationship exists in the target language.</td>
</tr>
<tr>
<td>• <strong>L1 Word Form</strong></td>
</tr>
<tr>
<td>Knowledge of formal (orthographic or phonetic) similarity between the target word or a part of it and a L1 word.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Non-Linguistic Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>• <strong>World Knowledge</strong></td>
</tr>
<tr>
<td>Non-linguistic knowledge, including knowledge of the topic of the text and other related background knowledge.</td>
</tr>
</tbody>
</table>

**Table 2** The Wesche & Paribakht Model Categorization (adapted from Wesche and Paribakht, 2009)

Inferencing is an internal activity and is often difficult to recognize. There are broadly two approaches to collect data when analyzing inferencing skills. First is offline
data, which focuses on how readers inferred implicitly. Second is online data, which
readers infer explicitly. The answers participants give in response to questions about a
text after reading the whole text are considered offline data, because the data is collected
after comprehension. In the case when the participants answer questions about the text
while comprehending a text, such as think-aloud tasks, the data is considered online data,
because the data is collected during comprehension. (Li and D’Angelo, 2016)

Wesche and Paribakht (2009) suggested that lexical inferencing is trainable, and
that it will give readers the confidence to deal with unfamiliar words. They also explained
in what ways lexical inferencing can be successful for L2 readers, and how L2 readers
can improve their lexical inferencing skill. First, they suggested that the reading material
should be a topic that is familiar and engaging for the L2 reader. Second, the text should
be comprehensible on their own. Hu and Nation (2000) discussed that when 98% of the
words contained in the text is understandable by the L2 reader, the reader can infer the
target word meaning more accurately. Third, thematic reading can provide readers to
encounter target words many times and can offer them more information about the target
word. Fourth, the readers should be advised to verify if the inference they made is
correct. Since L2 readers have a high possibility of not inferring accurately, they suggest
the readers use dictionaries and confirm after comprehending. Fifth, providing readers
knowledge on how to perform lexical inference is helpful. The knowledge includes
informing the readers what cue types and knowledge sources are available when
performing lexical inference. Finally, and most importantly, it is crucial to practice
frequently. Wesche and Paribakht mentioned that like other processes related to reading,
lexical inferencing also takes time and effort to develop and polish. Even if the readers
have an understanding of how they can infer word meanings in the text, it cannot be performed well without practice. In sum, readers can develop lexical inferencing when they use time to practice reading as much as they can.
CHAPTER 3
LITERATURE REVIEW

3.1 Prior Research on Inference in Reading

3.1.1 A Study on ESL Learners in Canada

Nassaji (2003) conducted research on how intermediate-level ESL learners perform in inferring word meanings from context clues in a reading a text. Twenty-one adult ESL learners participated in the study. The participants had different L1 backgrounds; 2 Arabic, 8 Chinese, 6 Persian, 2 Portuguese, and 3 Spanish. The participants and the reading materials were carefully selected to match with their appropriate level. The study collected the data by the “think-aloud procedure.” The participants were introduced to this procedure and were trained to understand how to verbalize their thoughts. They were also asked to infer target words by verbalizing and reporting whatever came to their mind. After finishing reading, the participants had time to review the passage and make additional comments about any of their thinking processes.

The result of this research indicated that the ESL learners were not very successful at inferring word meanings from context clues in the text. The study used a 3-point scale (2=successful, 1=partially successful, 0=unsuccessful) to rate their inference to the target words. Of the total inferential responses, 25.6% were successful, 18.6% were partially successful, and 55.8% were unsuccessful. This result indicated that students’ inferences were unsuccessful for more than half of the time. When looking at their
inference to individual words, the successful inferencing ranged from 9.5% to 38.1%. The study also pointed out that when unfamiliar words surround the target word, they tended to infer unsuccessfully. These results show that the words used in the text should meet the 98% familiarity, as mentioned in the previous chapter.

Furthermore, the readers seemed to infer the words from how they looked, and thus had confusion with similar-looking words. Although there were times when inferring the target word by how it looked was successful, this research showed that these items were either unsuccessfully or partially successfully inferred for 83.35% of the time on average. This finding suggests that what the word looks like and how it is similar to other unrelated words is a problem in inferring word meanings. It also indicated how accurately learners should recognize the word to infer accurately.

The study also investigated what knowledge sources the participants used to infer the target words. This research used the Nassaji model for categorization of knowledge sources. As a result, the participants used world knowledge most frequently (46.2%), followed by morphological knowledge (26.9%), grammatical knowledge (11.5%), discourse knowledge (8.7%), and L1 knowledge (6.7%). These numbers indicated the participants’ heavy reliance on their general knowledge when inferring the given text.

3.1.2 A Study on ESL Learners in Japan

Suzuki (2016) looked at how elementary level ESL learners perform in inferring words by looking at accuracy, use of knowledge sources, and use of reading strategies. The participants were 72 Japanese university freshmen who majored in international
business studies. They were all Japanese L1 students. The data was analyzed using data obtained from 60 participants who completed all the questions for this research. Their ages were 18 or 19 years old, and each had over six years of experience in learning English at educational institutions in Japan. The text which was used in the study contained 229 words and 13 target words. The coverage of familiar words was 94.3%.

After conducting the research, some of the target words were eliminated because they turned out to be familiar words to the participants. Therefore, there were 7 target words for analysis. The lexical inferencing accuracy was 14.7% on average, minimum 11.5% and maximum 18.8%. However, this result may be an overestimation because the target words cannot be considered unknown by all the participants when considering their background in English education. The participants may have had some knowledge in the target words and used their knowledge when inferring.

After reading the given text, the participants answered a questionnaire. The questionnaire contained questions to find out what kind of knowledge source the participants used to infer the target words. The Wesche & Paribakht model was used; however, the punctuation knowledge and text style/register knowledge were excluded from the taxonomy because it was unnecessary for the participants in this study. As a result, the participants did not use knowledge sources very much and could not make full use of it to infer the meanings of the target words. However, word knowledge was more used compared with other knowledge sources. Therefore, the study pointed out that the participants had a tendency in using knowledge sources included in the target words itself and suggested the possibility of knowledge source as a useful process if the reader has exact knowledge in it.
3.1.3 A Study on Japanese Language Learners

Yamagata (2013) did a study on 56 students who study Japanese as an L2 and who attended a higher-education institution in Japan. 41 students were Korean native speakers, and 15 students were Chinese native speakers. The students’ level was near-advanced intermediate to advanced level, and they were asked to answer all the questions on the survey and interviews in Japanese. There were four texts used in this research: two texts in which the topic was familiar for the readers, and two that were less familiar.

In order to see the participants’ lexical inferencing accuracy, the fill-in-the-blank test was prepared. The target words were selected by the necessity in the text, and were: 4 nouns, 4 verbs, and 2 adjectives. Although there was a model answer, whether their answers are correct or wrong was decided by if it fits the context well or not. Their answers were graded by a 3-point scale (2=successful, 1=partially successful, 0=unsuccessful). The result on how accurate the participants inferred word meanings was about 50%: 54.4% in texts which the topic is familiar to the participants and 50.6% in texts which the topic is less familiar. What was found was that there was not any significant difference in the accuracy to infer word meanings from a text that is familiar than one which is less familiar.

The study conducted a 10 minutes interview with each participant to find out what kind of knowledge source the participants used to perform lexical inferencing. The participants were asked to talk freely about the cues they used, how they came up to the answer, how they felt during the process of inferring meanings of unknown words, and what other possible answers they thought about during the process of comprehension.
After the interview, their answers were divided into four knowledge source categories depending on their answers. The study referred to both the Wesche & Paribakht model, as well as the Nassaji model. The four categories were: sentence knowledge, discourse knowledge, knowledge about languages other than the target language, and world knowledge. The most frequently used knowledge source was sentence knowledge, which was three times larger than the second most used knowledge source, which was discourse knowledge. The study also pointed out that the participants used about 26.5 knowledge sources to infer 20 target words. This result suggests that there are times when readers use various knowledge sources to comprehend a word.

3.2 Extensive Reading

Extensive reading is a method of foreign language learning by reading large amounts of texts in the target language. Since the mid-1980s, various linguists have been proponents of the inclusion of the extensive reading method in foreign language education. The paper Fukumoto (2004) shows that Harold Palmer founded the term extensive reading as an English language education approach. Harold Palmer argued that while intensive reading focuses on understanding the text line by line, extensive reading instead emphasizes understanding the overall meaning of the text, and reading at a faster pace. Below is the chart originally advocated by Waring (2009) and summarized by the writer that shows how the intensive reading method and extensive reading method is at odds with one another.
Table 3 Comparison of Intensive Reading and Extensive Reading

<table>
<thead>
<tr>
<th>Intensive Reading</th>
<th>Extensive Reading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language focus</td>
<td>Why?</td>
</tr>
<tr>
<td>Very little</td>
<td>Amount?</td>
</tr>
<tr>
<td>Hard</td>
<td>Difficulty?</td>
</tr>
<tr>
<td>Teacher</td>
<td>Who selects?</td>
</tr>
<tr>
<td>Text books</td>
<td>What?</td>
</tr>
<tr>
<td>In class</td>
<td>Where?</td>
</tr>
<tr>
<td>With exercises</td>
<td>Comprehension check?</td>
</tr>
<tr>
<td>Understand the whole text</td>
<td>Understanding?</td>
</tr>
</tbody>
</table>

There are four general rules in Japanese extensive reading which extensive readers should follow (Awano et al., 2012):

1. Learners should start reading from easy texts
2. Learners are not allowed to use dictionaries
3. Learners must skip over parts that they do not understand
4. Learners should quit reading when the story is not enjoyable

According to a summary by Day & Bamford (1998), there are various skills which the learners will acquire when they increase the amount of reading in the target language. One of the skills readers can acquire through reading massive amounts of texts in the target language is reading fluency. As said above, the readers should choose materials that they can read smoothly; thus, extensive reading focuses on learners to read fluently in addition to understanding the overall idea of the reading material. To accomplish fluent reading, knowledge which helps readers to comprehend texts is
necessary. Day & Bamford (1998) claims that extensive reading can give students a chance to develop knowledge that is necessary in reading such as linguistic, topical, and world knowledge.

Furthermore, one of the primary rules of extensive reading that is not familiar with language learners is the *non-use* of dictionaries. As mentioned above, extensive reading discourages readers from using dictionaries. The reason for this is that readers should choose reading materials that they can smoothly read and understand without getting much help from dictionaries. Further, by restricting the use of dictionaries, the readers will learn how they should deal with a word they do not know. Dealing with unknown words on their own is a technique that readers should acquire to accomplish fluent reading. The techniques include guessing or ignoring unknown words. According to de Bot et al.’s research (1997) research, 80% of the readers tried to infer the meaning of unknown words in a text and uses various knowledge sources.

In Yoshimura’s study (2018), the participants answered that through extensive reading, they learned a way to figure out if the reading material is a level which they could read without using dictionaries or not. Moreover, when comparing the responses obtained at the start of the semester and the end of the semester, the participants were comfortable in reading without using dictionaries in a more difficult level at the end of the semester. These responses from participants indicate that because the readers learned to read without using dictionaries through extensive reading, the readers are also comfortable in dealing with and inferring the meanings of unknown words. Day & Bamford (1998) argues that guessing or ignoring unknown words is a technique which readers of second languages should acquire. Because extensive readers have more
experience in reading and encountering unknown words, they have more experience in inferring unknown words and should have better inferencing skills than non-extensive readers.

3.3 Significance of This Study

This study examines whether the experience of extensive reading provides any influences on the ability to guess word meanings. There are three reasons why this study should be done. Firstly, no study has looked at the correlation of lexical inferencing ability and extensive reading even though extensive reading is one of the ways to acquire better word guessing skills. As it is “extensive” reading, language learners who have experience in extensive reading have more adventures in reading L2 texts than non-extensive readers. Extensive readers can have more reading practice through extensive reading, and so, extensive readers can build more reading comprehension ability than non-extensive readers. As noted in the previous chapter, inferencing is one of the cognitive processes used to comprehend a text. If extensive readers have more experience in reading and have better comprehension ability, the readers also have better ability in inferring unknown word meanings. Also, extensive readers have less stress in reading L2 texts because they are already used to handling reading in L2. When the readers are used to reading an L2 text as a whole, the ability and quality of lexical inferencing may also be better than those who have anxiety in reading L2 texts.

Second, extensive readers have more experience in encountering unknown words. As mentioned in the earlier section, extensive reading suggests the readers not to use dictionaries while practicing extensive reading. When readers cannot understand a
word and cannot comprehend a word, the advisors recommend guessing the meaning from contextual clues; thus, extensive readers have encountered to a situation which they have to infer word meanings more than non-extensive readers. Therefore, extensive readers have further experience in inferring word meanings. Moreover, when readers have extensive practice in inferring word meanings, the readers should also have more knowledge sources that are useful in lexical inferencing. As said in the earlier chapter, the use and quality of knowledge source differ from the readers’ background in education. The readers’ experience in extensive reading is one example of different education from non-extensive readers; therefore, how extensive readers use the knowledge sources may be different from non-extensive readers.

Lastly, there are not many studies which have focused on elementary and intermediate level L2 learners’ word guessing skills. This is understandable because elementary or intermediate level learners often do not have much experience in reading. Because of this, to conduct research is very difficult. However, extensive readers who are elementary or intermediate level in their language levels do have experience in reading L2 texts, at least a lot more than L2 learners at the same level. Therefore, there may be some differences between elementary and intermediate level extensive readers and non-extensive readers’ ability to infer word meanings.

3.4 Research Questions

This thesis will address the following two research questions from the data obtained from students who study Japanese as an L2 at University of Massachusetts Amherst.
1. Do extensive readers infer the meaning of unknown words more accurately than non-extensive readers?

2. Are there any differences in the use of knowledge sources to infer unknown words between extensive readers and non-extensive readers?
CHAPTER 4

METHODS AND PARTICIPANTS

4.1 About the Participants of This Research

The total number of participants for this study was 41 participants in total, while 39 participants completed all the questions on this research. The data from 39 participants was used for analysis in this study. Among the 39 participants, 13 students were enrolled in the introductory level Japanese language courses, 22 students were enrolled in the intermediate level Japanese language course, and 4 students were advanced level students. Among the participants, 11 participants were enrolled in the extensive reading course at the time of data collection or had enrolled and experienced extensive reading in the past semester; 1 student at the elementary level, 8 students at the intermediate level, and 2 students at the advanced level. Among the 39 participants, 38 students belonged to the University of Massachusetts Amherst, and 1 student belonged to a private university in the same area; however, the student who attended a private university was taking both the language course and the extensive reading course with the other 38 students at University of Massachusetts Amherst.

4.2 Research Design

This study is divided into two parts; survey section (see appendix A) and reading section (see appendix B). The survey asks participants about their reading habits such as how much they read in a week both in their L1 and Japanese, their experience in
extensive reading, how much they like to read, and so on. The survey was created to understand the participants’ background in reading and their general thoughts towards reading.

For the reading section, two reading materials were prepared to collect data. The first reading material is titled “Yūdachi.” This material is in a textbook for 1st graders used in some of the elementary schools in Japan (Mitsumuratosho, 2011a). The text is about a bunny and a raccoon who accidentally see each other after their fight they had the day before, but the sudden evening shower (yūdachi in Japanese) help them to reconcile and become friends again. The second reading material is titled “Amedama” by Nankichi Nīmi which is from a 5th graders’ textbook used in elementary schools in Japan (Mitsumuratosho, 2011b). The text is about a woman, the woman’s two children, and a samurai that ride together on a boat, but when the woman is having trouble with her two children begging for a candy drop (amedama in Japanese), the samurai cuts the candy into half using his sword.

The first reading material, Yūdachi, is written with only a few kanji and used simple sentences. The second reading material, Amedama, uses more complex structures and included more kanji. Table 4 and 5 show what and how many morphemes are used in the first 8 sentences in the two reading materials. When comparing the two, the second reading material, Amedama, included more information in a sentence because there were 16.25 morphemes in a sentence on average in Amedama whereas there were 12 morphemes in a sentence on average in Yūdachi. Further, Amedama had a higher possibility of using various types of morphemes in a sentence. For example, the first reading material, Yūdachi, did not include adjectives and determiners in the first 8
sentences; however, *Amedama* included both adjectives and determiners. From these differences, *Amedama* included more information in a sentence.

<table>
<thead>
<tr>
<th></th>
<th>Noun</th>
<th>Adjective</th>
<th>Verb</th>
<th>Particle</th>
<th>Adverb</th>
<th>Interjection</th>
<th>Auxiliary verb</th>
<th>Determiner</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total</strong></td>
<td>25</td>
<td>0</td>
<td>14</td>
<td>31</td>
<td>3</td>
<td>3</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td>3.13</td>
<td>0</td>
<td>1.75</td>
<td>3.88</td>
<td>0.38</td>
<td>0.38</td>
<td>2.5</td>
<td>0</td>
</tr>
<tr>
<td>Morphemes used in the first 8 sentences</td>
<td>Total = 96</td>
<td>Average = 12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4 Morphemes used in the First Reading Material “Yūdachi”

<table>
<thead>
<tr>
<th></th>
<th>Noun</th>
<th>Adjective</th>
<th>Verb</th>
<th>Particle</th>
<th>Adverb</th>
<th>Interjection</th>
<th>Auxiliary verb</th>
<th>Determiner</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total</strong></td>
<td>28</td>
<td>5</td>
<td>20</td>
<td>45</td>
<td>5</td>
<td>2</td>
<td>23</td>
<td>2</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td>3.5</td>
<td>0.63</td>
<td>2.5</td>
<td>5.63</td>
<td>0.63</td>
<td>0.25</td>
<td>2.88</td>
<td>0.25</td>
</tr>
<tr>
<td>Morphemes used in the first 8 sentences</td>
<td>Total = 130</td>
<td>Average = 16.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5 Morphemes used in the Second Reading Material “Amedama”

The reading materials provided readings of kanji by putting *furigana* on every kanji. Regarding the materials’ use in school textbooks, the 1\textsuperscript{st} reading material was expected to be easier than the 2\textsuperscript{nd} reading material; however, according to informal conversations with the participants, almost all the participants answered that the second reading material was easier to understand. This suggests that the difficulty level for native speakers and language learners differs, the way they read to comprehend texts differs as well.

Each reading material consisted of 6 target words; thus, there were 12 target words which the participants were asked to infer the meaning of using context clues. The target words were 3 nouns and 3 verbs in each reading material with a total of 6 nouns and 6 verbs. All the selected target words were then changed into pseudowords. Because the participants varied in their language levels from elementary to advanced, changing the target words into pseudowords that nobody had ever seen was the best way to control
the novelty of the target words. The Esperanto language was referred to create all the pseudowords in the reading material. It was decided to use the Esperanto language to generate the target words because the Esperanto language is an artificial language and it will be consistent to replace the target words referring to a language that is already created. Table 6 and 7 show the original word in the text, the equivalent in Esperanto language, and the pseudowords that were created for this study. As we can see, the equivalent of the target word in Esperanto language was changed into a form that suits the Japanese text.

<table>
<thead>
<tr>
<th><strong>Yūdachi</strong></th>
<th>Original</th>
<th>Meaning in English</th>
<th>Esperanto</th>
<th>Pseudo word</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>あめ</td>
<td>Rain</td>
<td>pluvo</td>
<td>ぶるうお</td>
</tr>
<tr>
<td>2</td>
<td>おさえました</td>
<td>Closed</td>
<td>reteni</td>
<td>れてにました</td>
</tr>
<tr>
<td>3</td>
<td>むきました</td>
<td>Turned to</td>
<td>direktīgi</td>
<td>ちれました</td>
</tr>
<tr>
<td>4</td>
<td>かみなり</td>
<td>Thunder</td>
<td>tondro</td>
<td>とんどろ</td>
</tr>
<tr>
<td>5</td>
<td>よりそって</td>
<td>Stuck together</td>
<td>karesumi</td>
<td>かれすみて</td>
</tr>
<tr>
<td>6</td>
<td>て</td>
<td>Hand</td>
<td>mano</td>
<td>まの</td>
</tr>
</tbody>
</table>

Table 6 List of Pseudowords and How It was Determined (1st reading material)

<table>
<thead>
<tr>
<th><strong>Amedama</strong></th>
<th>Original</th>
<th>Meaning in English</th>
<th>Esperanto</th>
<th>Pseudo word</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>のりました</td>
<td>rode</td>
<td>rajdi</td>
<td>らよていました</td>
</tr>
<tr>
<td>2</td>
<td>ひげ</td>
<td>beard</td>
<td>barba</td>
<td>ばるぼ</td>
</tr>
<tr>
<td>3</td>
<td>さしだしました</td>
<td>held out</td>
<td>etendi</td>
<td>あんたました</td>
</tr>
<tr>
<td>4</td>
<td>こまって</td>
<td>in trouble</td>
<td>embarasiği</td>
<td>えもぱって</td>
</tr>
<tr>
<td>5</td>
<td>刀</td>
<td>sword</td>
<td>glavo</td>
<td>ごらうお</td>
</tr>
<tr>
<td>6</td>
<td>もと</td>
<td>previous</td>
<td>juano</td>
<td>じゅあ</td>
</tr>
</tbody>
</table>

Table 7 List of Pseudowords and How It was Determined (2nd reading material)

The twelve pseudowords were all placed into brackets in the reading material which the
participants were asked to read. This was to confirm that the participants understand that the word is a target word that they were being asked to infer.

In the reading material, five selected words were underlined around each target words. Among the five words, two words were context clues that were helpful or necessary to accurately infer the target word. These underlined words were used in multiple-choice questions to see which word the participants used to infer each of the target words.

In the reading section, there was also an inference task that the participants were asked to complete. There were three questions:

1. What is the meaning of this word in English? (Open-ended question)
2. Select TWO clues that are strong to guess the meaning of the target word. (Multiple-choice question)
3. How did you guess the meaning? (Open-ended question)

The participants answered these three questions for each of the target words. Before obtaining data from the participants, three Japanese native speakers answered the above questions as pilot data. The pilot data were collected to understand what clues native speakers use, and what are the available contextual clues in order to infer the target word meaning. The five underlined words were determined based on the pilot data.

4.3 Data Collection

An Institutional Review Board (IRB) approval was obtained for human subject protection from the IRB office from University of Massachusetts Amherst (see appendix C). The permission to conduct this research was received on November 29th, 2018. The
data were collected from December 3rd, 2018 through March 8th, 2019.

One of the researcher’s fellow graduate students helped with the process of recruiting participants due to the instruction by IRB. The fellow graduate student visited language classes and extensive reading class during their class time to inform participants and explain about the process of this research, namely that it would take about 30 minutes and that they would be asked first to answer a questionnaire about their reading habits and then read texts and guess the meaning of the target words. Students who were willing to volunteer to participate in the research were asked to write down their names on the sign-up sheet. The students who signed up to participate were asked to come to Herter Hall in the University of Massachusetts Amherst at a specific time.

Before the process of collecting data, the participants first signed the consent forms. The participants had a chance to read the consent form before they signed it. The research consisted of three parts; a survey about their reading habits (see appendix A), questions regarding reading material 1, and questions regarding reading material 2 (see appendix B). The participants could choose between using the Google Survey form online or using a hard copy form to answer the survey.

After answering the first part which asks participants about their reading habits, reading material 1 and 2 were handed out. The participants were explained that the words in squares are the target words that are made up and the words do not exist in the Japanese language. The participants were also informed that two of the five underlined words around the squared words were useful words to help them find out what the target words mean, and asked them to select two underlined words that they thought were useful to infer the target words among the five underlined words. The participants also
were asked to write down what time they started reading and finished reading for each of
the reading materials.

4.4 Data Analysis

This research was conducted in order to answer the two research questions
below. Both the research questions were analyzed using the data collected from the
participants’ answers to the survey and from the comprehension questions in the reading
materials. The data obtained from the participants to answer the two research questions is
online data because the participants answered the questions during comprehension.

RQ1: Do extensive readers infer the meaning of unknown words more
accurately than non-extensive readers?

This research question was analyzed using the answers to the question “what is
the meaning of this word in English?” The participants were divided into two groups
depending on their responses in the survey: 1) participants who had no experience in
extensive reading and 2) participants who had experienced or were currently practicing
extensive reading. The answers were marked using a 3-point scale (2=successful,
1=partially successful, and 0=unsuccessful). Table 8 shows how the points were
determined in detail.
The points were summarized to show each individual's accuracy in inferring the target words. Their points were then calculated in both the extensive readers’ group and non-extensive readers’ group to determine if there were any differences in inferential accuracy between extensive readers and non-extensive readers.

**RQ2. Are there any differences in the use of knowledge sources to infer unknown words between extensive readers and non-extensive readers?**

This research question used the participants’ answers to the question in the reading section: How did you guess the meaning? This question was an open-ended question; therefore, the participants could answer based on what they thought and what context clues they used to infer the target words with their own words. To answer this research question, the participants’ answers were categorized into five knowledge sources: word knowledge, sentence knowledge, discourse knowledge, world knowledge, and L1 knowledge. This categorization refers to the Wesche & Paribakht model. The participants’ answers were analyzed to find out what knowledge sources each participant used to infer the target words and then summed up how many times each participant used...
which knowledge sources. Table 9 shows the taxonomy of five knowledge sources.

<table>
<thead>
<tr>
<th>Word</th>
<th>When the participants are using information of the target word itself.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentence</td>
<td>When the participants are using information in the immediate sentence.</td>
</tr>
<tr>
<td>Discourse</td>
<td>When the participants are using information outside of the immediate sentence or described the flow of the story.</td>
</tr>
<tr>
<td>World</td>
<td>When the participants are using information that they own.</td>
</tr>
<tr>
<td>L1</td>
<td>When the participants are borrowing information from their L1.</td>
</tr>
</tbody>
</table>

**Table 9 Taxonomy of Five Knowledge Sources**

The participants were divided into two groups: extensive readers’ group and non-extensive readers’ group. Two groups of students were compared to see if there were any specific knowledge sources that were used when inferring the target word.

The above taxonomy (Table 9) was also used to find if there were any differences between the two groups in the number of knowledge sources they utilized.

Some of the participants used more than one knowledge source types to infer a target word. For example, one participant answered the question for the first target word "ぷるゔぉ" saying, “Based on the sky and wind and the rabbit running under the tree.” The information, “wind and the rabbit running under the tree” was included in the immediate sentence, and the information “sky” was included in the previous sentence. The participant used information from both the immediate sentence; sentence knowledge, and the previous sentence; discourse knowledge. From this answer, we can see that the participant used both the sentence knowledge and the discourse knowledge to infer the target word “ぷるゔぉ.” Thus, this answer was counted as using 2 knowledge sources, because the participant used 2 types of knowledge sources (sentence knowledge and discourse knowledge) to infer one target word.
CHAPTER 5
ANALYSIS AND RESULTS

5.1 Analyzing Research Question 1

The first research question, “Do extensive readers infer the meaning of unknown words more accurately than non-extensive readers?,” was analyzed using the students’ answers to the first question for each of the pseudowords: what is the meaning of this word in English? There were 6 pseudowords in each reading; thus, there were 12 pseudowords in total. The method used in Nassaji’s (2003) and Yamagata’s (2013)’s research —rating the participants’ answers using a 3-point scale (2=successful, 1=partially successful, and 0=not accurate)— was used to measure participants’ answers. Below shows the percentage of successful inferencing in each language level.

<table>
<thead>
<tr>
<th></th>
<th>Beginner</th>
<th>Intermediate</th>
<th>Advanced</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Extensive Readers</td>
<td>10.42%</td>
<td>17.86%</td>
<td>20.83%</td>
<td>14.88%</td>
</tr>
<tr>
<td>Extensive Readers</td>
<td>25.00%</td>
<td>22.92%</td>
<td>79.17%</td>
<td>33.33%</td>
</tr>
</tbody>
</table>

Table 10 Inferencing Accuracy for the First Reading Material

<table>
<thead>
<tr>
<th></th>
<th>Beginner</th>
<th>Intermediate</th>
<th>Advanced</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Extensive Readers</td>
<td>22.92%</td>
<td>30.95%</td>
<td>33.33%</td>
<td>27.68%</td>
</tr>
<tr>
<td>Extensive Readers</td>
<td>33.33%</td>
<td>43.75%</td>
<td>66.67%</td>
<td>46.97%</td>
</tr>
</tbody>
</table>

Table 11 Inferencing Accuracy for the Second Reading Material

As seen in both Tables 10 and 11, extensive readers had higher percentages in answering the meanings of pseudowords correctly. An independent-samples t-test was conducted to compare intermediate and advanced level non-extensive readers and
extensive readers conditions. The elementary level participants were excluded because the number of elementary level extensive readers and non-extensive readers were too uneven to compare. For the first reading material, there was not a significant difference in the scores for intermediate and advanced non-extensive readers ($M=0.36$, $SD=0.40$) and intermediate and advanced extensive readers ($M=0.68$, $SD=0.65$) conditions; $t(13)=-1.33$, $p=0.206$. For the second reading material, there was also not a significant difference in the scores for intermediate and advanced non-extensive readers ($M=0.62$, $SD=0.39$) and intermediate and advanced extensive readers ($M=0.97$, $SD=0.47$) conditions; $t(16)=-1.83$, $p=0.086$.

Several students had a background in Chinese characters or kanji. Although the reading materials both included kanji, there were no differences in the accuracy of inferring word meanings between participants who have a background in kanji and who do not have a background in kanji.

5.2 Interpretation of the Analysis (RQ1)

As we can see from the analysis, there were no statistical differences in the accuracy of inferring unknown words among extensive readers and non-extensive readers. There are three possible reasons for not finding a significant difference among the two groups: 1) The reading material was too difficult to infer the target words accurately, 2) the amount of reading was not enough to see the effect of extensive reading to see the differences between extensive readers and non-extensive readers, and 3) there were not enough participants to compare the two groups.

Regarding the first possible reason, both Tables 10 and 11 show how all levels of
non-extensive readers, elementary level extensive readers, and intermediate level extensive readers’ inferring accuracy was relatively the same or below the results shown in the previous studies introduced in the previous chapter. However, advanced level extensive readers could infer more than half of the words successfully. This may be because the advanced level extensive readers who participated in this study could use other context clues accurately; therefore, their inferences were also more successful than the other groups. As it was mentioned in the previous chapters, one needs to understand 98% of the context in order to infer word meanings successfully. We can assume that the advanced level extensive readers tend to infer word meanings more successfully because they were able to understand more of what the context says. On the other hand, non-extensive readers at elementary, intermediate, and advanced level, and extensive readers at elementary and intermediate level had difficulties in accurately inferring the meaning of the target words because they could not understand what the context says as well as the advanced extensive readers.

Regarding the second possible reason, according to Yoshii’s summary (2016), it is reported that participants’ read 10,000 to 50,000 words in studies that showed positive effects of extensive reading. Furthermore, the period of practicing extensive reading that showed positive effects varied from 5 weeks to a year and a half; however, in the research that showed a positive effect in 5 weeks, the participants practiced extensive reading for 4 to 6 hours a day. As this summary shows, in order to succeed with the extensive reading method, one needs to practice extensive reading with a concerted effort. Extensive readers who participated in this study had experience in extensive reading for more than 12 weeks. The amount they are to read in a week is less than 50 minutes in class and
more than 50 minutes outside class; thus, the minimum amount of reading time may be about an hour and a half. Looking at the studies that showed positive effects of extensive reading, the length of time and the amount of reading was not enough for the extensive readers in this study to show that extensive readers have more accuracy in inferring unknown words. Also, it indicates how it likely takes time and effort to gain and improve inferencing accuracy skills through extensive reading.

Finally, for the third possible reason, there were 1 out of 13 in the elementary level, 8 out of 22 in the intermediate level, and 2 out of 2 in the advanced level who were extensive readers. Therefore, it was difficult to compare the two groups at each level because of the unevenness. If there were more participants of extensive readers that evens out the number of both extensive and non-extensive readers’ group, there is a possibility of the results showing extensive readers’ accuracy in inferring word meanings.

Although there were no statistical differences in the accuracy of inferring word meaning, we can assume that the extensive readers have a tendency to accurately inferring the word meaning. Table 10 and 11 show how extensive readers had a higher percentage of accurately inferring the pseudowords. Since extensive readers had a better inferencing accuracy than non-extensive readers in this study, it is conceivable that with more participants and evenness of the numbers of extensive readers and non-extensive readers, there is a possibility that we could see statistically significant differences.

5.3 Analyzing Research Question 2:

The second research question, “Are there any differences in the use of knowledge sources to infer unknown words between extensive readers and non-extensive
readers?,” was analyzed using the students’ answers to the question: how did you guess the meaning? The participants answered this question freely with their own words. Their responses were used to determine what knowledge sources the participants used to comprehend the pseudowords. Answers that did not provide enough information to decide which knowledge sources they used to comprehend was not used.

The result showed that the participants relied heavily on sentence knowledge and discourse knowledge to infer the target words. When the participant used the information in the immediate sentence; in other words, the sentence that contains the target word, it is considered that the participant used the sentence knowledge. For example, one participant answered “I think it’s weather related because it’s talking about the wind blowing and I think ぽたぽた (potapota) is the sound of water dripping” to explain how the participant inferred the first target word, “ぷるうぉ” meaning “rain” in the text. The information about the wind blowing and the word ぽたぽた (potapota) which is an onomatopoeia used to describe the sound of water dripping was in the immediate sentence. It is considered that the participant used the sentence knowledge to infer what ぷるうぉ means from this answer. When the participant used the information outside of the immediate sentence or the information which participants understand from the flow of the story, it is considered that the participant used the discourse knowledge. For example, one participant answered “If they became friends under the tree, then it makes sense for them to つなぐ their paws” to explain how the participant inferred the target word, “まの” which was used in a sentence, “まのをつないではしりだしました” meaning “they held hands and started to run.” The participant used the information from the flow of the story which was that the rabbit and the raccoon were in a fight, but they became friends
under the tree. It was considered that this participant used the discourse knowledge because the information was the flow of the story, which is information outside of the immediate sentence.

5.3.1 The Usage of Knowledge Sources to Infer Target Words in the 1st Reading Material (Yūdachi)

Figures 3 and 4 show the knowledge sources participants used to comprehend pseudowords in the 1st reading material. As we can see, there is a difference in the use of sentence knowledge and discourse knowledge between participants who have experienced and have not experienced extensive reading. An independent-samples t-test was conducted to compare non-extensive readers and extensive readers conditions. There was not a significant difference in the usage of sentence knowledge for non-extensive readers ($M=3.46$, $SD=1.21$) and extensive readers ($M=3.45$, $SD=1.16$) conditions; $t(19)=0.02$, $p=0.982$. However, there was a significant difference in the usage of discourse knowledge for non-extensive readers ($M=1.89$, $SD=1.52$) and extensive readers ($M=3.36$, $SD=1.77$) conditions; $t(16)=-2.33$, $p<0.05$. 
Figure 3 Non-Extensive Readers- Percentages of Knowledge Sources Used to Infer Target Words (1st Reading Material)

Figure 4 Extensive Readers- Percentages of Knowledge Sources Used to Infer Target Words (1st Reading Material)
5.3.2 The Usage of Knowledge Sources to Infer Target Words in the 2nd Reading Material (*Amedama*)

Figures 5 and 6 show the knowledge sources participants used to comprehend the 2nd reading material. There was less difference in non-extensive readers and the extensive readers compared with the 1st reading material. An independent-samples t-test was conducted to compare non-extensive readers and extensive readers conditions. There was not a significant difference in the usage of sentence knowledge for non-extensive readers (\(M=4.46, SD=1.50\)) and extensive readers (\(M=5.00, SD=0.95\)) conditions; \(t(28)=-1.28, p=0.210\). There was also not a significant difference in the usage of discourse knowledge for non-extensive readers (\(M=2.29, SD=1.89\)) and extensive readers (\(M=2.91, SD=1.38\)) conditions; \(t(24)=-1.10, p=0.283\).

![Figure 5 Non-Extensive Readers- Percentages of Knowledge Sources Used to Infer Target Words (2nd Reading Material)](image)

- word
- sentence
- discourse
- world
- L1
5.3.3 The Amount of Knowledge Sources Utilized

This study also analyzed whether there was a difference between non-extensive readers and extensive readers in the number of knowledge sources participants utilized to infer target words. The data obtained from both 1\textsuperscript{st} and 2\textsuperscript{nd} reading materials were combined and analyzed. The result showed that there is a difference between non-extensive readers and extensive readers in the amount of knowledge sources participants used to comprehend the pseudowords included in the two reading materials. To comprehend the 12 target words, non-extensive readers used 13.32 knowledge sources on average whereas extensive readers used 17.09 knowledge sources on average. An independent-samples t-test was conducted to compare non-extensive readers and extensive readers conditions. There was a significant difference in the scores for non-extensive readers ($M=13.32$, $SD=4.34$) and extensive readers ($M=17.09$, $SD=4.36$) conditions; $t(18)=-2.34$, $p<0.05$. 

![Figure 6 Extensive Readers- Percentages of Knowledge Sources Used to Infer Target Words (2\textsuperscript{nd} Reading Material)](image)
5.4 Interpretation of the Analysis (RQ2)

As mentioned in the analysis, there was a significant difference in the use of discourse knowledge among non-extensive readers and extensive readers in the 1st reading material, but not in the 2nd reading material. Thus, looking at the difference between the 1st and the 2nd reading material is essential. The 1st reading material was meant to be easier to understand because it used easier vocabulary and simple sentences; however, in terms of inferencing, the 1st reading material was more difficult than the 2nd reading material. This was because the sentence length in the 1st reading material was too short. When sentences are too short, the inference process may require additional information outside of the immediate sentence that includes the target word. The difference of sentence structures used in the 1st reading material and the 2nd reading material may be the reason why there was a difference in the use of discourse knowledge among non-extensive readers and extensive readers in the 1st reading material. The 1st reading material did not have enough information in the immediate sentence which included the target word. When extensive readers encountered the 1st reading material, they implicitly understood that the information they can earn from the immediate sentence is limited; therefore, they tend to use discourse knowledge to gather information outside of the immediate sentence to infer the target words. In contrast, non-extensive readers relied heavily on sentence knowledge in spite of the fact that the information provided is limited. This result shows that extensive readers had better inferencing skills because they were able to choose to use the appropriate knowledge source among all the other available knowledge sources to infer word meanings.
By understanding the reasons why the 1st reading material showed a difference in the use of knowledge sources between extensive readers and non-extensive readers, we can understand that there are positive effects of extensive reading in gaining inferencing ability. Extensive reading is a practice in which readers read as much as they can and understand the overall idea of the texts. Through extensive reading, the learners acquired skills to enjoy and understand the flow of the texts and got used to reading without using dictionaries. As a result, extensive readers could search for information outside of the immediate sentence when the information in the immediate sentence is not enough to infer the target words. This is because they know how to focus on the information in the entire text. Further, extensive readers knew that many context clues and knowledge sources are available when encountering and dealing with unknown words. Therefore, they could immediately change to using discourse knowledge or combine two or more knowledge sources rather than relying on just the sentence knowledge when there is not enough information in the immediate text. From these results, we can assume that extensive reading enhanced learners to acquire inferencing skills because they were able to choose the appropriate knowledge source when inferring the target words.

Moreover, extensive readers used a higher number of knowledge sources to infer word meanings. As it was said in Yamagata (2013), readers can use more than one knowledge sources to infer the meaning of a word. We could see the same result from this study; however, what should be pointed out is the number of knowledge sources used was different among extensive readers and non-extensive readers. The t-test also showed that the difference in the number of knowledge sources used between the two groups was significant. This result indicated that extensive readers have more available knowledge
sources through further inferencing experience than non-extensive readers. Extensive readers are more fluent in choosing and using knowledge sources when inferring word meanings. These discussions indicate that extensive reading can give tips to learners on how to apply knowledge sources when inferring word meanings and can improve inferencing skills.
6.1 Limits of This Study

There are several things to point out on regarding this study in the hope of making improvement when conducting similar studies. First, this study used Google Forms or hard copy forms for the participants to answer questions. However, many studies conducted interviews or use a think-aloud process to figure out what knowledge source the readers’ used for inference. Although it was possible to figure out the knowledge source used from words written by the participants, there were times when the participants did not write enough information to determine their knowledge source. Therefore, conducting an interview or using a think-aloud process may have helped to determine participants’ intention in using particular knowledge sources as well as their inference processes.

Secondly, it is important to note that it may have been possible to find more differences between non-extensive readers and extensive readers if there were more participants who had experienced extensive reading. Although this was a quantitative study, due to the uneven balance of extensive readers and non-extensive readers’ number at various levels, it was difficult to compare the two groups. If there were more extensive readers in each level and a greater number of participants in the total study, comparing students within the same language levels with more answers would be possible. Further, if there were more participants at each level, it would also be possible to see the participants’ accuracy in inferring word meanings by creating reading materials that
confirm 98% familiarity for each language level.

6.2 Conclusion

This study examined whether extensive reading gives positive effects to language learners’ reading comprehension, especially in their inferencing skills. There are many abilities which the learners must acquire in order to infer word meanings accurately. This study focused on the accuracy and the use of various knowledge sources when inferring unknown words. There were 39 participants, of which 11 were learners who had experience in extensive reading. The participants varied in their language levels, but were divided into three levels: an elementary level, an intermediate level, and an advanced level. The data obtained from the participants was used for analysis. The participants were mainly split into two groups for analysis: extensive readers and non-extensive readers. This two groups were created to find out if there were any differences in the accuracy and the use of knowledge sources among extensive readers and non-extensive readers.

There were no significant differences in the accuracy of inferring word meanings between the two groups. This may be because the reading materials that were used in this study did not match the 98% familiarity of the text. Further, the relative lack of reading amount of extensive readers and the unevenness of the number of participants between the two groups was also pointed out as one of the reasons why there was not much difference between extensive readers and non-extensive readers in terms of inferencing accuracy. Although there were no statistical differences, the accuracy percentage showed that the extensive readers’ group were more accurate in inferring word meanings. If there
were more participants, it is highly considered that we could find a difference in lexical inferencing accuracy between non-extensive readers and extensive readers. From this result, we can assume that extensive readers have a higher possibility of accurately inferring the word meaning than non-extensive readers.

The second focus of this study; the use of knowledge sources showed an interesting result. The result confirmed that there is a difference in the use of knowledge sources among extensive readers and non-extensive readers. Although non-extensive readers relied heavily on sentence knowledge to infer meaning, extensive readers were able to use discourse knowledge when there was not much information in the immediate sentence to infer the meaning of the target words. Further, extensive readers utilized more knowledge sources compared to non-extensive readers. This result shows that extensive readers have more choice of knowledge sources to draw from depending on the information they need and that they can also combine knowledge sources to infer word meanings. It should be noted that the extensive readers’ group were neither informed about nor trained on the usage of various knowledge sources. Through practicing extensive reading, they implicitly learned the ways to infer word meanings and were able to choose appropriate knowledge sources. As mentioned in the previous chapter, Wesche and Paribakht (2009) suggested that lexical inferencing is trainable through reading practice. This study confirms that lexical inferencing is trainable through extensive reading because extensive readers showed the ability to choose appropriate knowledge sources.

In summary, the results of this research showed how extensive reading has the potential in improving inferencing skills, and in a broader sense, can develop reading
comprehension skills in a foreign or second language context. As is mentioned in this study, reading comprehension consists of many various skills and inferencing skills is one of the many skills used in reading comprehension. Extensive reading should have more potential benefits for learners in gaining reading comprehension skills; therefore, extensive reading should be analyzed more.
APPENDICES
APPENDIX A

SURVEY SECTION OF THE RESEARCH

Questionnaire
This questionnaire asks you about your background in Japanese language and your reading habits. Please answer the following questions.

• What is your name? (Your name will be anonymous after collecting this data.)

• Including prior self study and middle school/high school classes, how long have you studied Japanese?

• Which Japanese language course are you currently enrolled in? If you are not enrolled in a Japanese course this semester, please choose the most recent Japanese language course you were enrolled in.
  ○ 110
  ○ 120
  ○ 126
  ○ 246
  ○ 326
  ○ 327
  ○ 497
  ○ 597

• Do you often read Japanese books (including e-books), manga, papers, newspapers, magazines, etc. for pleasure on regular basis? (Other than your textbook or other assigned readings)
  • Yes  
  • No

• If yes, how many minutes/hours per week do you read? (including Tadoku class)
  ○ I answered “no” to the previous question
  ○ ~30 minutes
  ○ 30 minutes - an hour
  ○ 1-2 hours
  ○ 2-3 hours
  ○ 3-4 hours
  ○ 4-5 hours
  ○ 5-6 hours
  ○ 6 hours or more
• Have you ever taken the Extensive Reading class (Tadoku class)?

  Yes  ·  No

• If yes, how many semesters have you taken the Extensive Reading class?
  ○ I answered “no” to the previous question
  ○ one semester
  ○ two semesters
  ○ three semesters
  ○ four semesters or more

• Are you currently enrolled to the Extensive Reading class?

  Yes  ·  No

• What is/are your first language(s)?

• Do you often read books (including e-books), papers, magazines, newspaper, etc. for pleasure in your FIRST language on regular basis? (Other than your textbook or assigned readings)

  Yes  ·  No

• If yes, how many minutes/hours per week do you read in your FIRST language?

  ○ I answered “no” to the previous question
  ○ 30 minutes
  ○ 1-2 hours
  ○ 3-4 hours
  ○ 4-5 hours
  ○ 5-6 hours
  ○ 6 hours or more
· If you were asked to choose between a book or a movie remake of the book, which would you choose? The content is the same.
  ○ Book
  ○ Movie remake of the book

· Do you like to read? (in any languages)
  I don’t like to read.  1 / 2 / 3 / 4 / 5  I love to read.

· Which feeling below apply the most when you encounter 2-3 words you do not know on a page while reading a Japanese text?
  ○ I would be frustrated.
  ○ I would be confused.
  ○ I would be a little bothered.
  ○ I would be curious about the word.
  ○ I wouldn’t feel anything. I just read on.

· What do you think the best reading strategy is when there are 2-3 words you do not know on a page while reading a Japanese text?
  ○ It is best to skip/ignore the word and read on to finish reading the text faster.
  ○ It is best to look the word up immediately on a dictionary/internet.
  ○ It is best to guess the meaning from the neighboring sentences/words.

· What would you do if you encountered 2-3 words on a page you did not know while reading in your FIRST language?
  ○ I am more comfortable when looking the word up on a dictionary/internet.
  ○ I like guessing meanings from the neighboring sentences/words.
  ○ I tend to skip/ignore the word and read on to read the text faster.
  ○ I tend to skip/ignore the word and read on because looking up on a dictionary is troublesome.

  ○ Other
• What would you do if you encountered 2-3 words on a page you did not know while reading a Japanese text?
  ○ I am more comfortable when looking the word up on a dictionary/internet.
  ○ I like guessing meanings from the neighboring sentences/words.
  ○ I tend to skip/ignore the word and read on to read the text faster.
  ○ I tend to skip/ignore the word and read on because looking up on a dictionary is troublesome.
  ○ I haven’t had enough opportunities to read Japanese texts to answer this question.
  ○ Other

• There are two Japanese books in front of you. You are choosing one to read “for pleasure”. Which book would you choose?
  ○ A book that is relatively difficult for your language level but the content is interesting to you.
  ○ A book that fits your language level but the content is not quite interesting to you.

Questions about the reading material will follow from the next page.
Please answer what time it is right now when you start. After answering the time, please start reading 「ゆうだち」 and answer the following questions. The words asked here are all words that do not exist in Japanese. There is no time limit but please try to read as fluently as possible. When you are done, please write down the time you finished.
ゆうだち

そこで、きょうに まっくらに なりました。小野と した かぜが ふき、
ばたばたと ぶるぶるが おちで きました。
「ゆうだちだ。」

くさばらに いた うさぎの こは、
おおいそで、水の したに かけこみました。

そこへ、
「ひゃあ、ずぶぬれだ。」
しぶきを ちらしながら、
たぬきの こが、とびこんで きました。

「やあ。」

と、いよいよ出ましたが、うさぎの こは、
あわてて 釣を れてにました。

たぬきの こも、
うさぎの こに きが つくと、
ぶいと ろくを られました。

にひきは、
きのう、けんかを したのです。
ぶりぶりして わかれたきり、
いちども つを きいて いません。

にひきは、
あいての かおを みないように して、
すこし はなれて たちました。
ぶるずれは、
さあさあと はげしく なり、
目も あけて られないほどです。

その とき、
にひきの あたまの うえで、
たたきつけるような おとが なりわりました。
どんなに です。にひきは、たおれるように
じめじめに ふせました。

それから すこし
じかんが たちました。

きが つくと、
にひきは、ぴったり
かれすみて いました。

そして、また しばらく たちました。

ぶるずれが あがり、あたりは、からりと
あかるく なりました。
「やんだ。」
「やんだ。」
にひきは、木の しかから とびだすと、
まだを つないで はしりだしました。

ゆうひが かっと でりつけ、
せみの こえが、きこえて きました。
What time is it now? （ ）

「ゆうだち」
・ぷるうお

What is the meaning of this word in English?

Select TWO clues that are strong to guess the meaning of ぷるうお.
☐ ひやり
☐ ばたばた
☐ ゆうだち
☐ くさはら
☐ うさぎ

How did you guess the meaning?

・れてにました

What is the meaning of this word in English?

Select TWO clues that are strong to guess the meaning of れてにました.
☐ 「やあ。」
☐ いいはじめましたか
☐ うさぎのこ
☐ あわてて口を
☐ たぬきのこ

How did you guess the meaning?
・ざれました

What is the meaning of this word in English?

Select TWO clues that are strong to guess the meaning of ざれました.

- きがつくと
- ふい
- よこを
- にひき
- きのう

How did you guess the meaning?

・どんどろ

What is the meaning of this word in English?

Select TWO clues that are strong to guess the meaning of どんどろ.

- そのとき
- あたまのうえで
- たきつけるようなおと
- たおれるように
- じめにふらせました

How did you guess the meaning?
What is the meaning of this word in English?

Select TWO clues that are strong to guess the meaning of かれすみて.

☐ きつつくと
☐ にひき
☐ びったり
☐ いいたれ
☐ しばらく

How did you guess the meaning?

What is the meaning of this word in English?

Select TWO clues that are strong to guess the meaning of まの。

☐ にひき
☐ 木
☐ したから
☐ つないので
☐ はじりだしました

How did you guess the meaning?

What time is it now? (  )
あめ玉

春のあたたかい日のこと、わたし席に二人の小さな子どもを連れた女の人から話していいました。

席が出ようとして、

「おうい、ちょっと待ってくれ。」

と、手の向こうから手をふりながら、さむらいが一人足ってきて、妻に飛びこみました。

席は出ました。

さむらいは妻の真ん中にどっかりすわっていました。ばかばかあたたかいので、そのうちにおむすびを食べました。

黒い立ぼを生やして張そうなさるかいが、こっくりこっくりするので、子どもたちはおかしくて、ふふふ、と笑いました。

お母さんは手に指を当てて、

「だまっておいで」

と言いました。さむらいが怒ったら、大変だからです。

子どもたちはだまりました。

しばらくすると、一人の子どもが、

「お駄け、あめ玉ちょうだい。」

と、手をあんたました。

すると、もう一人の子どもも、

「お駄け、あたしにも。」

と言いました。

お母さんは、ふところから紙のくるを取り出しました。ところが、あめ玉は、もう一つしかありませんでした。

「あたしにちょうだい。」

「あたしにちょうだい。」
二人の子どもは、強くお願いしました。お母さんは一つしかないので、お母さんはえもぱってまいりました。

「いい子たちだから、待っておいて。向こうへ着いたら、買ってあげるからね。」
と言って聞いてても、子どもたちは、

「ちょうどいいよう、ちょうどいいよう。」
と答えてくれました。

いねむりをしていたはずのさむらいは、ぱっちり目を覚まして、子どもたちがせがむのを見ていました。
お母さんはおどろきました。いねむりをじゃまされたので、このおさむらいは怒っているのにちがいない、
と思いました。

「おとしなくしておいで。」
と、お母さんは子どもたちをなだめました。

けれど、子どもたちは聞いてませんでした。

すると、さむらいがすすりと「ごらんお」と呟いて、お母さんと子どもたちの前にやって来ました。

お母さんは真っ青になって、子どもたちをかばいました。いねむりのじゃまをした子どもたちを、さむらい
が切ってしままったと思えたのです。

「おめえを出せ。」
と、さむらいは言いました。

お母さんは、おそおそとおめえを差し出しました。

さむらいはそれを寄のへりにのせ、ごらんおでばちんと二つに切りました。

そして、

「そうね。」
と、二人の子どもに向かってやりました。

それから、またじゅうの所に帰って、こっくりこっくりねむり始めました。
What time is it now? ( )

「あめ玉」
らよでいました
What is the meaning of this word in English?

Select TWO clues that are strong to guess the meaning of らよでいました.

☐ 春
☐ あたたかい日
☐ わたし舟
☐ 「ちょっと待ってくれ。」
☐ 土手の向こう

How did you guess the meaning?

・ぱるぱ
What is the meaning of this word in English?

Select TWO clues that are strong to guess the meaning of ぱるぱ.

☐ あたたかい
☐ ねむり始めました
☐ 黒い
☐ 生やして
☐ こっくりこっくり

How did you guess the meaning?
・あんたました
What is the meaning of this word in English?

Select TWO clues that are strong to guess the meaning of あんたました.
- 母ちゃん
- ちょうだい
- 手を
- もう一人
- 子ども

How did you guess the meaning?

・えもばって
What is the meaning of this word in English?

Select TWO clues that are strong to guess the meaning of えもばって.
- 強くお願いしました
- 一つしかない
- いい子たちだから
- 待っておいて
- 買ってあげる

How did you guess the meaning?
・ごらづお
What is the meaning of this word in English?

Select TWO clues that are strong to guess the meaning of ごらづお.

☐ 子どもたち
☐ すらり
☐ ぬいて
☐ お母さん
☐ やって来ました

How did you guess the meaning?

・じゅあ
What is the meaning of this word in English?

Select TWO clues that are strong to guess the meaning of じゅあ.

☐ 「そうれ。」
☐ 二人の子ども
☐ また
☐ 帰って
☐ こっくりこっくり

How did you guess the meaning?

What time is it now? (   )
APPENDIX C

IRB

Consent Form for Participation in a Research Study
University of Massachusetts Amherst

Researcher(s): Sayako Niwa (MA, Japanese)
               sniwa@umass.edu
               Yuki Yoshimura (Senior Lecturer, Japanese)
               yuki@umass.edu

Study Title: The ability to guess the meaning when reading L2 texts

1. WHAT IS THIS FORM?
This form is called a Consent Form. It will give you information about the study, so you can make an informed
decision about participation in this research. This consent form will give you the information you will need to
understand why this study is being done and why you are being invited to participate. It will also describe what
you will need to do to participate and any known risks, inconveniences or discomforts that you may have while
participating. We encourage you to take some time to think this over and ask questions now and at any other time.
If you decide to participate, you will be asked to sign this form and you will be given a copy for your records.

2. WHO IS ELIGIBLE TO PARTICIPATE?
Any students who have background in studying Japanese as a second language and is over 18 years old.

3. WHAT IS THE PURPOSE OF THIS STUDY?
The purpose of this research study is to find out language learners' abilities in guessing the meaning when reading L2
texts.

4. WHERE WILL THE STUDY TAKE PLACE AND HOW LONG WILL IT LAST?
The study will be conducted at W.E.B. Du Bois Library or Herter Hall at UMass Amherst campus and will take
about 30 minutes. No more participation will be needed.

5. WHAT WILL I BE ASKED TO DO?
There are two parts to the research study. In the first part you will be asked about your general reading habits and
your experience in extensive reading. In the second part, you will be asked to read a short story in Japanese that
includes pseudo words and will answer to a Google Form which asks what you think the pseudo words mean and
how you came up to that answer. You may skip any question you feel uncomfortable answering.

6. WHAT ARE MY BENEFITS OF BEING IN THIS STUDY?
You may not directly benefit from this research; however, we hope that your participation in the study may
connect to better language learning in SLA.

7. WHAT ARE MY RISKS OF BEING IN THIS STUDY?
A possible inconvenience may be the time it takes to complete the study. A data breach is possible; however, the
researcher has made every reasonable effort to maintain the confidentiality of the data.
8. HOW WILL MY PERSONAL INFORMATION BE PROTECTED?
The researchers will keep all study records, including any codes to your data, in a locked box stored in a locked room. Research records will be labeled with a code. A master key that links names and codes will be maintained in a separate and secure location. The master key will be destroyed 3 years after the close of the study. All electronic files such as databases and spreadsheets containing identifiable information will be password protected. Any computer hosting such files will also have password protection to prevent access by unauthorized users. Only the members of the research staff will have access to the passwords. At the conclusion of this study, the researchers may publish their findings. Information will be presented in summary format and you will not be identified in any publications or presentations.

10. WHAT IF I HAVE QUESTIONS?
Take as long as you like before you make a decision. We will be happy to answer any question you have about this study. If you have further questions about this project or if you have a research-related problem, you may contact the researcher, Sayako Niwa: phone number (413-345-3997). If you have any questions concerning your rights as a research subject, you may contact the University of Massachusetts Amherst Human Research Protection Office (HRPO) at (413) 545-3428 or humansubjects@ora.umass.edu.

11. CAN I STOP BEING IN THE STUDY?
You do not have to be in this study if you do not want to. If you agree to be in the study, but later change your mind, you may drop out at any time. There are no penalties or consequences of any kind if you decide that you do not want to participate.

12. WHAT IF I AM INJURED?
The University of Massachusetts does not have a program for compensating subjects for injury or complications related to human subjects research, but the study personnel will assist you in getting treatment.

13. SUBJECT STATEMENT OF VOLUNTARY CONSENT
When signing this form, I am agreeing to voluntarily enter this study. I have had a chance to read this consent form, and it was explained to me in a language which I use and understand. I have had the opportunity to ask questions and have received satisfactory answers. I understand that I can withdraw at any time. A copy of this signed Informed Consent Form has been given to me.

Participant Signature: ____________________________  Print Name: ____________________________  Date: ____________________________

By signing below, I indicate that the participant has read and, to the best of my knowledge, understands the details contained in this document and has been given a copy.

Signature of Person Obtaining Consent: ____________________________  Print Name: ____________________________  Date: ____________________________
REFERENCES


