English Reading Comprehension by Japanese High School Students: Structural Equation Modeling including Working Memory and L1 Literacy

Hideko Shibasaki1*, Shingo Tokimoto2, Yuichi Ono3, Tsugio Inoue4, Katsuo Tamaoka5

1General Studies and Education Development, Nagaoka University of Technology, Nagaoka, Japan
2Foreign Language Studies, Mejiro University, Shinjuku, Japan
3Graduate School of Humanities and Social Sciences, University of Tsukuba, Tsukuba, Japan
4Cultural Studies, University of Kochi, Kochi, Japan
5Graduate School of Languages and Cultures, Nagoya University, Nagoya, Japan
Email: shibalea@vos.nagaokaut.ac.jp, ktamaoka@lang.nagoya-u.ac.jp

Received 9 September 2015; accepted 10 October 2015; published 13 October 2015

Copyright © 2015 by authors and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY).
http://creativecommons.org/licenses/by/4.0/

Abstract
The present study aims to build a model of second language (L2) reading, taking into account working memory capacity and first language (L1) literacy as variables within structural equation modeling (SEM). In this study, a total of 120 Japanese L1 high school students were given a series of tests: 1) English reading span tests, 2) English vocabulary size tests, 3) English grammar tests, 4) Japanese literacy tests, and 5) two kinds of English reading tests. One of the English reading tests was on a general topic, and the other required specific background knowledge to control for the effects of topic familiarity. SEM revealed that L2 working memory, L2 grammatical knowledge, and L2 vocabulary knowledge significantly predicted general L2 reading, while general L2 reading, along with L1 literacy, was the significant predictor of L2 reading on a specific topic. Furthermore, L1 literacy predicted L2 vocabulary knowledge, and skilled L1 readers excelled at creating situation model for L2 reading comprehension, even when they lacked topical knowledge. The present study presents a more refined model of L2 reading comprehension, which accounts for L1 literacy and working memory more comprehensively than previous research has.

Keywords
L2 English Reading Comprehension, Working Memory, L1 Japanese Literacy, Structural Equation

*Corresponding author.

1. Introduction

The goal of this study is to create a more refined model of second language (L2) reading that includes working memory and first language (L1) literacy as variables within a structural equation modeling (SEM) framework. L2 reading is a highly complex endeavor, involving multiple cognitive processes that are supported by subskills, including vocabulary knowledge (Moghadam, Zainal, & Ghaderpour, 2012), background knowledge (Hauptman, 2000), visual word recognition (Hamada & Koda, 2011), phonological awareness (Zhang & Koda, 2013), syntactic processing (Bowden, Steinhauer, Sanz, & Ullman, 2013), inference (Alptekin, 2006), and prediction (Keung & Ho, 2009). These factors have been included in L2 reading models in previous studies. However, no L2 reading model includes working memory as a variable. In addition, many previous studies have investigated the relationship between L1 and L2 reading. Therefore, the present study aims to build a more refined model of L2 reading that includes working memory and L1 literacy as variables in SEM.

1.1. Verbal Working Memory

In L1 reading research, as a reading factor, working memory has been identified as one of the significant explanatory variables through a meta-analysis of 67 studies (Daneman & Merikle, 1996). Readers start by visually recognizing letters and then recognize words based on letter combinations. This process allows readers to identify meaning from their mental lexicon. In addition, syntactic knowledge enables them to understand the relationships among words, sentences, and paragraphs. Finally, readers determine the text’s cohesive meaning by integrating the text’s surface meaning with information retrieved from long-term memory. All of this processing takes place within working memory (Just & Carpenter, 1992).

Researchers have hypothesized that, when an individual reads text, he or she acquires and assimilates new information and simultaneously retains already processed information (Just & Carpenter, 1992). Such internal computation requires working memory (Baddeley & Hitch, 1974). Additionally, reading span tests (RSTs) have been used as an effective means of testing individual differences in this type of processing capacity (e.g., Cain, Oakhill, & Bryant, 2000, 2004; Daneman & Carpenter, 1983; Seigneuric & Ehrlich, 2005; Seigneuric, Ehrlich, Oakhill, & Yuill, 2000).

In their research on the role of working memory in L2 reading comprehension, Miyake and Friedman (1998) found that working memory is more strongly associated with syntactic comprehension in the L2 than in the L1. Other studies have suggested that L2 working memory is more closely associated with L2 reading comprehension than L1 working memory is (Harrington & Sawyer, 1992; Walter, 2004). Because of its prominent role in L2 reading, working memory was added as a necessary variable in our model.

1.2. Relationship between L1 Literacy and L2 Reading Ability

Many researchers have argued that L1 literacy (or L1 reading ability) and L2 reading comprehension have some things in common. Recent experimental studies on neural activity in L1 and L2 processing have also suggested that these factors are closely connected. According to Van Hell and Tokowicz (2010), eleven event-related potential (ERP) studies on syntactic processing found that L2 learners presented (biphasic) ERPs similar to L1 speakers when the L1 and L2 were structurally similar or when the L2 proficiency was high.

Various theoretical orientations have been discussed regarding the connection between the L1 and the L2, including the interdependence hypothesis, the threshold hypothesis, and the dual-language system. In addition, many studies have found that the working memory capacity in the L1 and L2 are strongly correlated (e.g., Osaka & Osaka, 1992; Osaka, Osaka, & Groner, 1993; Shibasaki, Tokimoto, Ono, & Inoue, 2015). As a result of these findings, L1 literacy has been included in the L2 reading model.

L2 reading research in the 1980s and the 1990s often entertained the notion that having relevant background knowledge would facilitate reading comprehension (e.g., Carrell, 1987; Floyd & Carrell, 1987; Hauptman, 2000; Hudson, 1982; Johnson, 1981; Taglieber, Johnson, & Yarbrough, 1988; Tudor, 1986, 1988). Shibasaki (2005) indicated that background knowledge was a significant predictor of situation model construction, which sug-
gested that having relevant knowledge would facilitate reading comprehension, just as in L1 reading (e.g., Takahashi, 1996). However, in daily life, the reader does not always have background knowledge of the texts being read. In fact, the reader is often challenged by having to learn or gain new knowledge from reading the text with little or no background knowledge of the subject matter. However, although individual differences exist, all L2 learners have acquired L1 literacy. Thus, by controlling for large individual differences in the background knowledge of a specific field, it was possible to focus on how L1 literacy contributed to the construction of a situation model in L2 reading and to the acquisition of knowledge from an L2 text.

1.3. Various Models of L2 Reading Comprehension

Various models exist in the field of L2 reading comprehension. Bernhardt (2000, 2005) reviewed the field of L2 reading from the 1970s to 2000 and presented a compensatory model of L2 reading that showed the relationship between L2 proficiency and L2 reading. Additionally, McNeil (2012) elaborated on Bernhardt’s (2005) model and added the strategic variables of knowledge and background knowledge. In McNeil’s model, the relative contributions of the variables shifted, depending on the reader’s L2 proficiency.

The models above are only speculative, based on findings from previous research. Although such speculative models are sometimes hypothesized as precursors to building path models from multiple regression or SEM (e.g., Shibasaki, 2005; Takahashi, 1996), the practical value of such models will continue to decrease until they are empirically tested. Moreover, the relationships between the variables are not shown in the compensatory models. Modeling a complex phenomenon such as L2 reading should entail not only identifying the variables but also specifying the relationships among the variables.

In an effort to further explore relationships among variables in L2 reading, Shibasaki (2005) sampled 120 Japanese high school students and presented a path model of L2 English reading comprehension. The dependent variables were two levels of representation (i.e., textbase and situation model), and the predictor variables were L2 linguistic knowledge, L2 vocabulary knowledge, and background knowledge. Textbase and situation model are different levels of memory representation in the theory of van Dijk and Kintsch (1983). The first level of memory representation in reading comprehension is the surface structure, which includes the actual words and phrases in the text. The second level of the representation is called textbase, which involves the surface structure of texts, i.e., sentence-level local coherence and propositions. A situation model is a representation that results from the integration of propositions from textbase and the knowledge that the reader has retrieved from long-term memory. According to the path model in Shibasaki (2005), L2 vocabulary knowledge and L2 linguistic knowledge had a strong correlation, and these two variables robustly predicted the L2 textbase. Moreover, the two levels of text representations were found to predict each other bidirectionally. Background knowledge did not predict textbase, but it was a significant predictor of situation model. This path model provided some insight into a question that attracted attention during the 1980s through the 1990s i.e., whether background knowledge or vocabulary knowledge more strongly affected L2 reading comprehension. The Shibasaki model showed that background knowledge and vocabulary knowledge contributed to different levels of reading comprehension.

However, although the path model can show the relationships between variables and is superior to the compensatory models of Bernhardt (2000, 2005) and McNeil (2012), it cannot help determine goodness of fit and also lacks generalizability, which would require the construction of a model based on SEM. The benefits of SEM are as follows. First, both latent and observed variables can be analyzed with SEM. Second, representing the complex relationships between the variables is possible with SEM. Third, evaluating the model’s goodness of fit is possible with SEM.

L2 reading models based on SEM analysis have been conducted by Guo and Roehrig (2011) and Zhang (2012). Guo and Roehrig (2011) constructed a model based on SEM, using 278 Chinese college students learning English. In their model, 87% of the variance in L2 reading comprehension was explained by L2 general reading knowledge and L2 specific knowledge.

Zhang (2012) constructed a model of L2 reading with data from 190 advanced learners of English whose first language was Chinese. In his model, the latent variables were vocabulary knowledge, grammatical knowledge, and L2 reading. The breadth and depth of the learner’s vocabularies were the observed variables for vocabulary knowledge, whereas grammatical judgment and grammatical error correction were the observed variables for grammatical knowledge. Zhang’s model indicated that as much as 81.1% of the variance in L2 reading could be predicted by both vocabulary knowledge and grammatical knowledge. Not only was the contribution of vocabu-
lary and grammatical knowledge congruent with findings from previous research, but Zhang’s model also showed a more than adequate goodness of fit \( N = 172, \chi^2(11) = 12.437, p = .332, \text{ns.} \) \( \text{CFI} = .990, \text{RMSEA} = .028. \). Therefore, we decided to use SEM to construct our model for L2 reading comprehension, including the observed variables of vocabulary knowledge, grammatical knowledge, working memory, and L1 literacy.

Determining the appropriate participants is one of the major considerations in a study such as ours. In addition to conducting an SEM analysis, Gottardo and Mueller (2009) tested 131 Spanish-speaking English learners in grade 1 (and later in grade 2) on L1 and L2 skills. Their results suggested that L2 phonological awareness significantly predicted L2 word reading and that L2 oral language proficiency and L2 word reading significantly predicted L2 reading comprehension. However, L1 oral language proficiency and phonological awareness were not significant predictors of L2 reading. For example, the influence of L1 literacy had been observed only for participants who had higher L2 proficiency, not for learners with lower L2 proficiency (Bossers, 1991; Hayashi, 2003; Pichette, Segalowitz, & Connors, 2003). Thus, the young participants’ incomplete and ongoing development of L1 proficiency was posited as the reason for the lacking contribution of L1 proficiency to L2 reading. These findings showed that, sampling older participants and participants with higher L2 proficiencies would be necessary to clarify the relationship between L1 literacy and L2 reading.

In Japan, most instruction in English as a foreign language starts only after students have finished primary education, by which time they have acquired their L1 literacy. Therefore, it was determined that the target participants would be high school students aged 15 and older.

2. Method

2.1. Participants

A total of 120 (97 male, 23 female) second-year students at a national engineering high school in Japan participated in the study. Their first language was Japanese. At the time of data collection, they had studied English for 4 years and 8 months, starting in the first year of middle school. The Global Test of English Communication (GTEC), developed by a Japanese company called Benesse, was used to measure the participants’ overall English proficiency. Of the GTEC tests, all participants took the Basic Type, consisting of reading, listening and writing, in December 2012. The mean total score of the three subsections (Max = 660) ranged from 237 to 583 points, with a mean score \( M = 389.48 \) and a standard deviation (SD) of 73.02. The overall test results revealed significant variance among the participants. For each section, the scores were Max = 250, \( M = 145.03, SD = 33.53 \) for reading; Max = 250, \( M = 151.38, SD = 35.54 \) for listening; and Max = 160, \( M = 102.07, SD = 20.22 \) for writing.

2.2. Tasks and Procedure

2.2.1. English Reading Span Test (RST)

The English RST (Shibasaki et al., 2015) was used for the study. This is the only test of its kind developed specifically for Japanese high school students, and it showed a relatively higher reliability (Cronbach’s alpha = .875). The English RST was administered in a lab using a computer-assisted language learning (CALL) system with individual distributions and recording capabilities for audio files via a headset. The participants were instructed to wear a noise-sealing, over-the-ear headset with a microphone. After the study’s objective was explained to the participants, the task instructions were provided.

The task had four conditions that ranged from two to five sentences, with five trials in each condition. The participants read aloud English sentences that appeared on the screen of the individually assigned computers. After one trial, the screen went blank, and the participants were instructed to repeat the last word of each sentence that they had just read aloud. The order of reproduction was unrestricted, as long as the last word of the last sentence shown was not given first. To ensure the full understanding of the instructions, the experimenter demonstrated both good and bad examples, which were followed by a few practice rounds. After the practice rounds, the CALL system played a PowerPoint presentation showing stimulus sentences for the English RST. The task began with the experimenter’s cue. Each participant’s voice was automatically recorded and converted into mp3 files. A total of approximately 2 GB of data were collected.

The experimenter listened to the recordings and assigned scores. The RST scoring methods include the proportion of words, the truncated span, correct word sets, and memory capacity. Memory capacity was used be-
cause it yields scores that are intervals and is capable of estimating the amount of information that can be stored at one time. The normality of the score distribution was checked using the Kolmogorov-Smirnov test. The result was not significant ($D = .909, p < .71, ns$), indicating that the assumption of normality was met.

### 2.2.2. English Vocabulary Size Test

Previous studies (e.g., Grabe, 2009; Laufer, 1992, 2000; Perfetti, Landi, & Oakhill, 2005) indicated that L2 vocabulary knowledge greatly contributes to L2 reading comprehension. Two dimensions of L2 vocabulary knowledge were conceptualized as the breadth and depth of vocabulary knowledge. More recently, Zhang (2012) showed that the standardized partial regression coefficient was greater for the breadth of vocabulary knowledge ($\beta = .86$) than for the depth of vocabulary knowledge ($\beta = .60$). Based on Zhang’s (2012) results, the breadth of vocabulary knowledge is included as a variable in the present study.

The Vocabulary Levels Test developed by Nation (1990) is frequently used to measure the size of one’s English vocabulary, but our pilot test indicated that this test was too difficult for Japanese second-year high school students. Therefore, instead of Nation’s test, another vocabulary test was used to measure size (Mochizuki, Aizawa, & Tono, 2003). This test was developed specifically for Japanese learners of English. Similar to the Vocabulary Levels Test by Nation (1990), this test measures one’s ability to choose the correct meaning of two words out of six choices. To make the test more accessible to Japanese high school students, the cue words are given in Japanese. For example, two Japanese words, *makige* (curly hair) and *niku/nikutai* (body), are given, and the participant has to choose the correct meaning from the following six choices: 1) beach, 2) curl, 3) economy, 4) flesh, 5) glory, and 6) worker. In this case, the correct answers are 2) for *makige* and 4) for *niku/nikutai*.

The target words for the replacement test were selected from the Japan Association of College English Teachers’ (JACET) 8000 Vocabulary. In this test, the number of words is assumed to reflect the participant’s vocabulary knowledge. In the JACET 8000, the highest vocabulary level is 8000 words for college-level English. Because the participants in this study are second-year Japanese high school students who are learning English as a second language, the present study used the 5000 word level. The mean score was 2821 ($SD = 402$), and scores ranged from 1633 to 3800.

### 2.2.3. English Grammatical Knowledge Test

Along with L2 vocabulary knowledge, L2 grammatical knowledge has also been identified as an important factor in L2 reading comprehension (e.g., Akbari, 2014; Bowey, 1986; Givón, 1995). Nassaji (2003) found a stronger correlation between L2 vocabulary knowledge and L2 reading comprehension ($r = .53, p < .001$) than between L2 grammatical knowledge and L2 reading comprehension ($r = .44, p < .001$). By contrast, Zhang (2012) revealed that the standard partial regression coefficient was larger for L2 grammatical knowledge ($\beta = .66, p < .05$) than for L2 vocabulary knowledge ($\beta = .42, p < .05$) in relation to L2 reading comprehension. As a result, both L2 vocabulary and grammatical knowledge are now seen to be major factors in L2 reading comprehension.

Zhang (2012) used two tasks to measure grammatical knowledge (i.e., the error correction task and the grammaticality judgment task). His study found that the grammaticality judgment task was weakly correlated with L2 reading comprehension ($r = .174, p < .05$), whereas the error correction task did not show a significant correlation. Therefore, for the current study, a grammaticality judgment task was adopted.

The grammaticality judgment task had 40 items, which included grammar rules presented in English textbooks used in middle schools and high schools in Japan. The grammar rules included tenses, direct/indirect speech, comparisons, subjunctives, infinitives, gerunds, participles, auxiliary verbs, passive voice, relative clauses, prepositions, and conjunctions. This task had multiple-choice items, with 4 choices per item. Counting each item as one point, the mean score was 18.47 points ($SD = 4.17$), and the scores ranged from 6 to 26.

### 2.2.4. Two Types of L2 Reading Tests

This study used two types of tests to measure L2 reading comprehension. In one of these tests, topics used in the texts are general and do not require specialized background knowledge. This generality contrasts with the content of the other test, in which the topics discussed in the text require specialized background knowledge of a specific field.

The L2 reading test for general knowledge (hereafter, L2 general reading) consisted of 36 multiple-choice
items, with 4 options each, which were selected from the 2012 version of the Basic GTEC for the second year of high school. The reasons for using this version of the test were as follows: 1) The GTEC is a large-scale test with high reliability, and 2) the topics used in the test do not require any specific background knowledge (the effect of the participants’ background knowledge thus being minimized).

The first task in L2 general reading is an item in which the reader has to search for the meeting place for a tour in a pamphlet or for specific information in a newsletter. The second task involves reading a text of approximately 500 words and choosing a sentence that adequately summarizes the text. The third task requires forming a sentence with word-, phrase-, and clause-level coherence. Finally, the fourth task involves reading a text of approximately 500 words and making inferences about causal relationships, explanations, the characters’ feelings, and the author’s intentions.

The scores in each of the aforementioned tasks represent the following observed variables for the latent variable of L2 general reading: 1) information search, 2) gist, 3) sentence structure, and 4) textual inference. The mean score, the score range, and the standard deviation of each test are shown in Table 1.

As for the L2 reading test for specialized background knowledge (hereafter, L2 specific reading), a reading comprehension test on tied aid was created as a task requiring background knowledge of a specific field (Appendix). This test was created after obtaining permission from the original creator of the text used in Kobayashi (2002).

Tied aid is a form of bilateral aid from a developed country to a developing country in which the procurement of the goods or services involved is limited to the donor country and often consequently makes the developing country more dependent on the donor country. The participants in this study, who were majoring in engineering, were assumed to not be familiar with topics such as economics and international relations.

When the participants were asked if they had known what “tied aid” was, only 17 of 120 said “yes,” and 113 said “no.” Of the 17 participants who answered “yes,” none was able to correctly explain what tied aid was when asked to do so in a follow-up question. These answers verified that the assumption that participants lacked knowledge of tied aid was correct. In turn, the text could be assumed to be a type of text from which building situation model was difficult because the participants were not able to make use of their background knowledge. The minimized effects of background knowledge in constructing situation model while reading was assumed to highlight how other variables affect L2 reading comprehension. Because the test elicits open-ended responses, a scoring rubric was created to ensure rigorous and systematic scoring.

The vocabulary and grammar used for the text were controlled to be only those introduced by the end of the third year of middle school. Keenan and Betjemann (2006) and Kintsch and Kinstch (2005) have found that some reading comprehension questions used in the previous experiments could be answered without reading the actual text. To prevent this issue, a pilot test was conducted with ten college students to see if they could correctly answer the questions without reading the text. For this task, none of the pilot test participants was able to answer the questions correctly without reading the text, which suggested that the text did not present this particular problem.

The four observed variables for the latent variable of L2 specific reading are as follows: 1) explicit information, 2) cohesion, 3) knowledge acquisition, and 4) paragraph integration. The mean score, the score range, and the standard deviation for each test are shown in Table 1.

### 2.2.5. Japanese Literacy Test

For the Japanese literacy test, 11 items on kanji (Chinese characters) and vocabulary and 17 items on reading comprehension were selected from the 2013 version of the Japanese Language Standard Achievement Test created by the Society of High School Teachers for the Study of the Japanese Language. All items were multiple-choice questions, with five choices per item. Japanese has four scripts (kanji, hiragana, katakana, and the Roman alphabet), along with a rich and diverse vocabulary with four types of words (i.e., kango, words of Chinese origin; wago, words of Japanese origin; loan words; and hybrid words). Consequently, vocabulary knowledge is an important aspect of literacy. Children are expected to learn 2126 kanji by the end of their compulsory education (i.e., the end of middle school), and being able to read and write kanji is a major factor in Japanese literacy.

The aforementioned test has an expository text and a narrative text. The questions involve resolving anaphora, summarizing, finding causal relations among events, and interpreting characters’ feelings and the author’s intentions. The mean score was 55.97 ($SD = 12.49$) out of 100, with a Cronbach’s alpha of .698.
Table 1. Descriptive statistics and correlations of all observed variables.

<table>
<thead>
<tr>
<th>#</th>
<th>Variables</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>L2 Working memory</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>L2 Vocabulary size</td>
<td>.18</td>
<td>.26</td>
<td>.26</td>
<td>.35</td>
<td>.35</td>
<td>.35</td>
<td>.35</td>
<td>.35</td>
<td>.35</td>
<td>.35</td>
<td>.35</td>
<td>.35</td>
</tr>
<tr>
<td>4</td>
<td>L2 Information search</td>
<td>.28</td>
<td>.19</td>
<td>.19</td>
<td>.35</td>
<td>.35</td>
<td>.35</td>
<td>.35</td>
<td>.35</td>
<td>.35</td>
<td>.35</td>
<td>.35</td>
<td>.35</td>
</tr>
<tr>
<td>5</td>
<td>L2 Gist</td>
<td>.29</td>
<td>.29</td>
<td>.29</td>
<td>.33</td>
<td>.33</td>
<td>.33</td>
<td>.33</td>
<td>.33</td>
<td>.33</td>
<td>.33</td>
<td>.33</td>
<td>.33</td>
</tr>
<tr>
<td>6</td>
<td>L2 Sentence structure</td>
<td>.30</td>
<td>.19</td>
<td>.19</td>
<td>.42</td>
<td>.42</td>
<td>.42</td>
<td>.42</td>
<td>.42</td>
<td>.42</td>
<td>.42</td>
<td>.42</td>
<td>.42</td>
</tr>
<tr>
<td>7</td>
<td>L2 Textual inference</td>
<td>.15</td>
<td>.27</td>
<td>.27</td>
<td>.47</td>
<td>.47</td>
<td>.47</td>
<td>.47</td>
<td>.47</td>
<td>.47</td>
<td>.47</td>
<td>.47</td>
<td>.47</td>
</tr>
<tr>
<td>8</td>
<td>L2 Explicit information</td>
<td>.15</td>
<td>.25</td>
<td>.25</td>
<td>.18</td>
<td>.18</td>
<td>.18</td>
<td>.18</td>
<td>.18</td>
<td>.18</td>
<td>.18</td>
<td>.18</td>
<td>.18</td>
</tr>
<tr>
<td>9</td>
<td>L2 Cohesion</td>
<td>.08</td>
<td>.09</td>
<td>.09</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>10</td>
<td>L2 Integration of contents</td>
<td>.29</td>
<td>.35</td>
<td>.35</td>
<td>.63</td>
<td>.63</td>
<td>.63</td>
<td>.63</td>
<td>.63</td>
<td>.63</td>
<td>.63</td>
<td>.63</td>
<td>.63</td>
</tr>
<tr>
<td>11</td>
<td>Acquisition of knowledge from L2 text</td>
<td>.24</td>
<td>.17</td>
<td>.17</td>
<td>.22</td>
<td>.22</td>
<td>.22</td>
<td>.22</td>
<td>.22</td>
<td>.22</td>
<td>.22</td>
<td>.22</td>
<td>.22</td>
</tr>
<tr>
<td>12</td>
<td>L1 literacy</td>
<td>.19</td>
<td>.17</td>
<td>.17</td>
<td>.40</td>
<td>.40</td>
<td>.40</td>
<td>.40</td>
<td>.40</td>
<td>.40</td>
<td>.40</td>
<td>.40</td>
<td>.40</td>
</tr>
</tbody>
</table>

Min 0.30  6.00  1633  0.00  0.00  0.00  0.00  0.00  0.00  0.00  23.00
Max 3.40  26.00  3800  8.00  5.00  11.00  10.00  5.00  6.00  5.00  90.00
Mean 1.82  18.48  2822  4.10  2.20  4.97  4.69  1.95  3.16  1.43  0.82  55.97
Standard deviation 0.51  4.18  402  1.81  1.30  2.47  1.95  1.96  1.64  1.57  1.12  12.95

Note: N = 91. *p < .05. **p < .01.

3. Results

From the initial pool of 120 participants, the final analyses were conducted on data from 91 participants by excluding those who did not complete all tasks and those who could not properly follow directions on the RST. Table 1 shows the descriptive statistics and the correlations between all the variables in the study. The SPSS Amos 22.0 (2013) package was used for the SEM analysis. The model was constructed with the reading test of a general text in English and the reading test of specialized text in English as the two latent variables. The other measured variables were the observed variables.

Table 2 shows the parameter estimates of all latent and observed variables. The results of the SEM model are depicted in Figure 1. The model yields an excellent goodness of fit [N = 91. x2(52) = 59.911, p = .211, ns, CFI = .954, RMSEA = .041].

The SEM model in Figure 1 consists of two latent variables and 12 observed variables. Each of the two latent variables of L2 general reading and L2 specific reading was constructed with four observed variables (eight in total). The remaining four observed variables were independently treated to investigate causal relations.

A confirmatory factor analysis for the latent variable of L2 general reading yielded significant standardized partial coefficients (β) for all four variables: sentence structure (β = .560, p < .001), information search (β = .547, p < .001), gist (β = .526, p < .001), and textual inference (β = .633, p < .001). All observed variables showed significant contributions to the latent variable. The confirmatory factor analysis for the latent variable of L2 specific reading also showed significant standardized partial coefficients for all of the observed variables, except for cohesion (β = .233, ns): explicit information (β = .717 p < .001); knowledge acquisition (β = .656, p < .001); and integration of contents (β = .514, p < .001).

The causal relation of L2 general reading to L2 specific reading was significant (β = .416, p < .001). Three observed variables were also significantly predicted in L2 general reading: L2 working memory (β = .299, p < .05); L2 grammatical knowledge (β = .250, p < .05); and L2 vocabulary knowledge (β = .588, p < .001). While both L2 knowledge of grammar and vocabulary contributed to L2 general reading, vocabulary knowledge was the stronger predictor for L2 general reading. L2 working memory significantly supported L2 general reading.
Figure 1. The model for the relationship between L2 general reading and L2 specific reading including L1 literacy and working memory.

Table 2. Parameter estimates of SEM for L2 text comprehensions.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Variables</th>
<th>β</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2 Vocabulary Size</td>
<td>&lt; --- L1 Literacy</td>
<td>.404</td>
<td>***</td>
</tr>
<tr>
<td>Reading Ability of L2 General Text</td>
<td>&lt; --- L2 Working Memory</td>
<td>.299</td>
<td>*</td>
</tr>
<tr>
<td>Reading Ability of L2 General Text</td>
<td>&lt; --- L2 Grammar Knowledge</td>
<td>.250</td>
<td>*</td>
</tr>
<tr>
<td>Reading Ability of L2 General Text</td>
<td>&lt; --- L2 Vocabulary Size</td>
<td>.588</td>
<td>***</td>
</tr>
<tr>
<td>Comprehension of L2 Text in Specific Field</td>
<td>&lt; --- Reading Ability of L2 General Text</td>
<td>.416</td>
<td>**</td>
</tr>
<tr>
<td>Comprehension of L2 Text in Specific Field</td>
<td>&lt; --- L1 Literacy</td>
<td>.373</td>
<td>**</td>
</tr>
<tr>
<td>Reading Ability of L2 General Text</td>
<td>&lt; --- L1 Literacy</td>
<td>.008</td>
<td>ns</td>
</tr>
<tr>
<td>L2 Sentence Structure</td>
<td>&lt; --- Reading Ability of L2 General Text</td>
<td>.560</td>
<td>***</td>
</tr>
<tr>
<td>L2 Information Search</td>
<td>&lt; --- Reading Ability of L2 General Text</td>
<td>.547</td>
<td>***</td>
</tr>
<tr>
<td>L2 Gist</td>
<td>&lt; --- Reading Ability of L2 General Text</td>
<td>.526</td>
<td>***</td>
</tr>
<tr>
<td>L2 Textual Inference</td>
<td>&lt; --- Reading Ability of L2 General Text</td>
<td>.633</td>
<td>***</td>
</tr>
<tr>
<td>L2 Explicit Information</td>
<td>&lt; --- Comprehension of L2 Text in Specific Field</td>
<td>.717</td>
<td>***</td>
</tr>
<tr>
<td>L2 Cohesion</td>
<td>&lt; --- Comprehension of L2 Text in Specific Field</td>
<td>.233</td>
<td>ns</td>
</tr>
<tr>
<td>L2 Integration of Contents</td>
<td>&lt; --- Comprehension of L2 Text in Specific Field</td>
<td>.514</td>
<td>***</td>
</tr>
<tr>
<td>Acquisition of Knowledge from L2 Text</td>
<td>&lt; --- Comprehension of L2 Text in Specific Field</td>
<td>.656</td>
<td>***</td>
</tr>
</tbody>
</table>

Note: N = 91. *p < .05. **p < .01. ***p < .001.

The observed variable of L1 literacy predicted both L2 vocabulary knowledge (β = .404, p < .001) and L2 specific reading (β = .373, p < .01). These two causal relations between the L1 and L2 indicated close interlanguage relations. Two subgroups were created based on their total L1 literacy scores to reveal the areas in L2 specific reading affected by L1 literacy. The high group consisted of the top 30 participants in L1 literacy (M = 66.9, SD = 10.3), and the low group consisted of the bottom 30 participants in L1 literacy (M = 45.6, SD = 10.9).
L2 reading comprehension of a specific field was predicted by four observed variables (i.e., explicit information, cohesion, knowledge acquisition, and integration of contents). As shown in Table 3, the t-tests revealed that the two groups only significantly differed in knowledge acquisition.

4. Discussion

4.1. Contribution of Working Memory to L2 Reading

Working memory has been found to be a significant predictor in the models of L1 reading comprehension (e.g., Leong, Tse, Loh, & Hau, 2008; Takahashi, 1996, 2001). The L2 reading model in the current study also found working memory to be a significant predictor of L2 reading comprehension. Proficient L2 readers with rich vocabulary knowledge can allocate many processing resources to higher-order processing because their lower-level processing, such as character and word recognition, is automatized and does not require many cognitive resources. On the other hand, unskilled L2 readers with inadequate linguistic knowledge tend to use up processing resources in lower-level processing and thus have few remaining resources for higher-order processing, such as referencing world knowledge to compare with the text or constructing cohesive meaning representations. In other words, the ability to efficiently use limited processing resources differentiates proficient readers from non-proficient readers. This distinction is more pronounced in L2 reading than in L1 reading.

The model in this study suggests that working memory predicts L2 reading comprehension better than L2 grammatical knowledge. In L1 reading research, Takahashi (2001) presented a model based on five-year longitudinal data on Japanese elementary school students. According to the model, as the students moved up in grades, the effect of working memory capacity on L1 reading comprehension decreased, and vocabulary knowledge became a strong predictor of L1 reading comprehension.

In the current study, participants were high school students who had studied English for 4 years and 8 months. As their English proficiency improves, the effect of their working memory capacity may change. This exciting possibility is something that could be addressed in another study. However, this study’s significance lies in its use of an English RST on Japanese high school students, which has not been employed in any previous studies.

4.2. The Contribution of L1 Literacy to L2 Reading

The current study found the following relationships between L1 literacy and L2 text comprehension: First, although L1 literacy was correlated with the information search, gist, and textual information variables (Table 1), it was not a significant predictor of L2 reading of general topics (Table 2, Figure 1); second, L1 literacy significantly predicted the L2 reading of a specific field text, where the use of background knowledge was restricted (Table 2, Figure 1); and, third, skilled readers with high L1 literacy were significantly better at learning (i.e., acquiring new knowledge) from reading an L2 text, even when they could not use their background knowledge (Table 3).

At the end of the experiment, the participants were instructed to circle any word in the texts that they did not understand.

Table 3. Differences on L2 text comprehension of the upper and lower group by L1 literacy.

<table>
<thead>
<tr>
<th>Variables of L2 reading specific field</th>
<th>Group</th>
<th>Mean</th>
<th>SD</th>
<th>t-value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explicit information</td>
<td>Upper</td>
<td>2.30</td>
<td>1.90</td>
<td>1.76</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td>Lower</td>
<td>1.43</td>
<td>1.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohesion</td>
<td>Upper</td>
<td>3.17</td>
<td>1.84</td>
<td>−0.08</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td>Lower</td>
<td>3.20</td>
<td>1.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acquisition of knowledge</td>
<td>Upper</td>
<td>1.10</td>
<td>1.09</td>
<td>2.32</td>
<td>**</td>
</tr>
<tr>
<td></td>
<td>Lower</td>
<td>0.50</td>
<td>0.90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Integration of contents</td>
<td>Upper</td>
<td>1.63</td>
<td>1.67</td>
<td>0.79</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td>Lower</td>
<td>1.30</td>
<td>1.60</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: N = 30 (upper group). N = 30 (lower group). **p < .01.
H. Shibasaki et al.

know. More than half of the participants did not know either “tied” (53.1%) or “aid” (67.3%). This finding indicates not only that the “tied aid” concept was unknown to all the participants but also that the component words of the compound word “tied aid” were also unknown to most participants.

How, then, did the participants manage to comprehend this text? For question 1 (i.e., “What do rich countries want to happen in poor countries?”), the participants had to find the following sentence within the same paragraph: “They hope that poor countries will no longer need aid in the future and can be independent.” The participants then had to determine that “they” referred to “rich countries.” In question 2, the reader had to connect the underlined “their” (1) with the plural noun “the other countries,” which preceded “their” in the same sentence, and then find its referent, “poor countries.” These two questions are connected to textbase and can be addressed with linguistic knowledge of English.

Question 4 reflected both textbase and situation model of the third paragraph. It required an understanding of the second, third, and fourth sentences of the third paragraph and their relationship to one another and then the integration of the information present in those sentences. For questions 1, 2, and 4, there were no significant differences based on L1 literacy performance (Table 3).

In contrast to the questions above, a difference in L1 literacy was shown in the students’ performance on question 3, which asked what tied aid was. This was a question that investigated how well the readers understood that which was unknown to them, i.e., how well they were able to acquire new knowledge from the text. The passage was about how tied aid hindered the independence of developing countries, even though, on the surface, tied aid seemed beneficial to these countries. On the rating scale, the highest score of 3 was awarded to those who were able to explain the harm caused by tied aid. When a concrete example of this harm, such as “The developing countries are forced to buy machinery and parts from the developed countries,” was mentioned, a score of 2 was awarded. Responses that merely included a reference to international relations or official assistance, without explicit mentioning the problems inherent in tied aid, received a score of 1.

Readers find drawing coherent meaning from an L2 text that includes unfamiliar vocabulary and background knowledge extremely difficult. However, in the aforementioned passage, careful readers should be able to infer that the topic relates to international relations or international aid because of its use of vocabulary such as “poor countries” and “rich countries.” These words are the first clues in understanding “tied aid,” but unskilled readers are said to have difficulty in detecting these first clues (Hampton & Resnick, 2009). The very ability to detect such clues may be attained by developing L1 literacy. After correctly inferring that the topic relates to international aid, understanding what “tied aid” is and what problems it might introduce will require understanding the relationships between the second and third paragraphs, after successfully integrating the information in these paragraphs.

Another subskill required in reading, which poor readers often lack, is the ability to form coherent meaning by connecting smaller units, such as words, phrases, and sentences (Hampton & Resnick, 2009). Skilled L1 readers can form meaning from information from these smaller units, and they are expected to be able to transfer this subskill to L2 reading. The text on tied aid had a text structure: first, it introduced a topic; second, it raised an issue; third, it provided a concrete example of this issue; and fourth, it revealed the author’s view on the issue. Readers with strong L1 literacy skills have been known to use their knowledge of text structures, as observed in previous studies (e.g., Koda, 1990; Kong, 2006; Salataci & Akyel, 2002). Therefore, in this study, readers with high L1 literacy may also have been able to take advantage of strategies developed from L1 literacy, including their knowledge of text structures to develop a situation model of tied aid, even though the L2 text presented an unfamiliar topic and vocabulary.

Understanding explicit information in the text or its coherence is often not the goal of reading a text. Instead, the goal may be to acquire new knowledge from the text. This study shows that L1 literacy plays an important role in reconstructing one’s knowledge using information taken from an L2 text. Studies have long claimed that L1 and L2 reading are related. However, the particular L1 subskills that influence particular L2 subskills have not been clarified. The SEM in this study elucidates the specific ways in which L1 and L2 literacy might be related.

4.3. The Contribution of L1 Literacy to L2 Vocabulary Building

L1 literacy was a strong predictor of L2 vocabulary knowledge in the current model ($\beta = .404$, $p < .001$). Although this finding was not expected and these two variables seem to be otherwise unrelated, L1 literacy may
actually play a major role in L2 vocabulary acquisition. Two reasons for these results are suggested below.

First, L1 vocabulary is a major supporting factor for L1 literacy. If an individual has a large L1 vocabulary, he or she likely has a large network of word concepts. L1 vocabulary acquisition is said to take place through labeling, packaging, and network building (Aitchison, 2003). Labeling involves making connections between the sequence of sounds and the entities that they represent. Packaging involves discovering the things that can be bundled together under the same label. Network building involves creating an understanding of how words relate to one another. By contrast, in L2 vocabulary acquisition, the learner often already knows both the concept and the label in their L1 (Mochizuki, Aizawa, & Tono, 2003). Hatch and Brown (1995) specified these processes as follows: “1) encountering new words, 2) getting the word form, 3) getting the word meaning, 4) consolidating word form and meaning in memory, and 5) using the word” (p. 374). In the third step, if the word concept in the L2 has not yet been learned in the L1, knowledge acquisition may prove quite difficult. Although the word concepts in the L1 and L2 do not necessarily match perfectly, learning the word without any concept of the word’s meaning would be impossible. If L2 vocabulary is acquired, to an extent, by assigning new labels to L1 vocabulary, having a rich stock of word concepts is inevitably necessary, which may explain why L1 literacy predicted L2 vocabulary size.

Second, the participants were not yet proficient in the L2 because they were only in their second year of high school—a stage where the target language was understood through translation into the L1, through which the effect of L1 literacy came into play. Because translation involves switching between the L1 and L2, the transition between the L1 and L2 does not work as smoothly without a solid L1 base. In fact, previous studies have found that L2 learners with developing proficiencies often replace the L2 with the L1 to understand an L2 text. For example, based on a study on inferences by proficient and non-proficient English learners whose L1 was Japanese, Muramoto (2000) asserted that proficient L2 learners directly understood the target language, whereas non-proficient L2 learners first translated the L2 sentence into their L1 to understand it. Muramoto’s idea is that proficient L2 learners can extract meaning directly from the L2 words, but non-proficient L2 learners will not know the meaning of the L2 words until they are translated into their L1. This assertion is congruent with the Revised Hierarchical Model (Kroll & Stewart, 1994), which claims that highly proficient L2 learners connect L2 vocabulary directly to concepts and that non-proficient L2 learners access concepts through the L1 vocabulary system.

The participants in the current study had studied English domestically for only 4 years and 8 months. Thus, they could be assumed to still understand English through its translation into Japanese. Additionally, the vocabulary size test asked participants to provide a Japanese translation for each English word, which might have had some effect. The results might have been different had this study been conducted with participants who had had higher levels of English proficiency.

5. Conclusion

This study tested 120 Japanese high school students on six tasks: an English RST, an English vocabulary size test, an English grammar test, a Japanese literacy test, and two different kinds of English reading tests on a general topic and a specialized topic. Based on the scores of these tests, a model of L2 reading was constructed using SEM. This model differs from the models in previous studies because working memory and L1 literacy are included as variables.

This study produced two important findings: First, working memory was found to be a significant predictor of L2 reading when working memory was measured by the English RST, which had not previously been tested on Japanese high school students. Second, the study determined how L1 literacy affected L2 reading. Specifically, L1 literacy was found to be a significant predictor of L2 reading comprehension on a specialized topic and of L2 vocabulary knowledge. L2 learners with high L1 literacy excelled in acquiring new knowledge from an L2 text and in constructing situation model, even when the text was on a specialized topic for which they could not use their background knowledge.

These results suggest how important L1 literacy is for L2 learners. L1 literacy is considered to be the basis for all learning, including L2 learning. In recent years, early English education has been strongly promoted in Japan, and some plans suggest starting English conversation classes in elementary schools. These plans aim to develop children’s communication skills in English. However, some researchers have cautioned against the negative ef-
fect of early English language education. If elementary schools include English classes in their curricula, Japanese classes, among others, will be forced to reduce their class time. Brumfit and Johnson (1979) argued that if children learned a foreign language before they mastered their native language, they might lose their ability to think in their native language. They insisted that foreign languages should be learned after children were equipped with the basic ability to think. As communication skills refer to one’s ability to clearly express his or her thoughts and feelings so that other people can understand him or her, children need to learn communication skills in their L1 before learning an L2. The results of the present study suggest the importance of L1 literacy for English educators in Japan.

Adopting van Dijk and Kintsch’s (1983) theory of comprehension with two levels of representation (i.e., textbase and situation model) for L2 reading comprehension, this study has elucidated the complex relationship between L1 literacy and L2 reading. In future studies, investigating the relationships between one’s L1 literacy, L2 proficiency and background knowledge of a specific field would be worthwhile.

Acknowledgements

This study has received support from Grants-in-Aid for the Scientific Research Program (B) “Multilateral research of L1 and L2 Reading Comprehension; 2011-2014” (Principal Investigator: Hideko Shibasaki).

References


Appendix

Read the following English passage and answer the questions below in Japanese.

It is said that (a) developed countries possess 78% of all wealth in the world. This means that the other countries have only 22% of the wealth, even though their (1) population is approximately 76% of the world’s total. Many rich developed countries give aid to (b) poor countries. The purpose of giving aid is to help poor countries improve their situation. They hope that poor countries will no longer need aid in the future and can be independent.

However, many people argue that giving aid to poor countries does more harm than good. One example is “tied aid” (2). For example, machinery is given to a poor country under certain conditions. It usually means that (c) the receiving country has to spend the money on what is produced in (d) the giving country. In other words, if a poor country is given expensive tractors, agricultural productivity in (e) the country may improve rapidly, but when the tractors break down, they will have to buy expensive spare parts from the giving country. Therefore, “untied aid” has been considered desirable since the 1980s.

Another problem of giving aid is that most aid is used in cities, and it makes people leave the countryside and move to cities. The countryside becomes empty and can no longer produce enough food for its people. At the same time, if cities become overcrowded, there will be many problems, such as housing shortages, unemployment, and poor health care facilities. As a result, they will become poorer and poorer. We should consider what (f) developing countries really need.

1) What do rich countries want to happen in poor countries?
2) Put a “T” if the word(s) refer(s) to the same entity as the underlined “their” (1). Put an “F” if it the word(s) do(es) not refer to the same entity.
   (  ) a. developed countries
   (  ) b. poor countries
   (  ) c. the receiving country
   (  ) d. the giving country
   (  ) e. country
   (  ) f. developing countries
3) What is “tied aid” (2)?
4) Why is it a problem if people move from the countryside to cities?