The Effects of English Learner's First Language (L1) Background Content Knowledge on Guessing Unknown Words

MARIKO SUGITA
Temple University, Japan

This study was an investigation of the effects that learners' first language (L1) background content knowledge has on guessing unknown or partially known words when reading a written text. Twenty-nine Japanese university students of English with similar English reading proficiency read a 1,387-word narrative text in one of three L1 background knowledge conditions: reading an L1 summary before reading the English text, reading the L1 summary before reading the English text and having it available while reading the text, and reading the English text without reading the L1 summary. The participants were tested for their word-form recognition of 33 target words and they took a bilingual multiple-choice recognition test and two questionnaires immediately after reading. There was no systematic difference in guessing unknown words among the groups, as guessing the meaning of the target words occurred successfully across all groups. The qualitative results indicated that the participants believed that the L1 summary helped them comprehend a general idea of the text and guess the meaning of unknown words.

Sugita, M. (2016). The Effects of English Learner's First Language (L1) Background Content Knowledge on Guessing Unknown Words *Journal of Extensive Reading* 4(1). http://jalt-publications.org/jer/

Keywords: EFL learners, background knowledge, guessing from context, inferring from context, unknown words, kanji.

Introduction

During the past three decades, researchers have investigated incidental learning through guessing from context in reading a written text both with first language (L1) and second language (L2) learners. They have provided empirical evidence on how guessing from context occurs, what factors affect guessing, how much learning occurs through guessing, whether inferred words are retained, and whether incidentally learned words are retained (e.g., Joe, 1998; Pitts, White, & Krashen, 1989; Pulido, 2007; Saragi, Nation, & Meister, 1978). Mondria and Wit-De Boer (1991), for example, found that guessability and retainability did not work in tandem. Extensive reading (ER) can help learners incidentally acquire lexical knowledge (Waring & Takaki, 2003) and provide "suitable conditions for un-

known words from context" (Wodinsky & Nation, 1988, p. 155).

Other researchers have investigated guessing from a strategic point of view and found that guessing is a useful and trainable strategy for vocabulary learning (Gu & Johnson, 1996; Williams, 1985), while yet others introduced strategies for guessing words from context (Clarke & Nation, 1980). Walters (2006) investigated three methods of teaching inferring meaning from context (strategy training, context clue instruction and practice and feedback with cloze exercises) to investigate which means was most effective. Walters concluded that "the strategy instruction resulted in improved ability to infer from context, and that both context clue and strategy instruction resulted in improved reading comprehension" (pp. 187-188).

Nation (2001) stated that incidental learning through guessing from context is essential for vocabulary learning, and Schmitt (2000) has regarded guessing as a key skill for vocabulary learning. According to Laufer (1997), readers' background knowledge of the subject matter of a text contributes to successful guessing in some cases, while in others this knowledge can mislead readers or suppress correct interpretations of the text. Few researchers, however, have examined what effects learners' L1 background content knowledge has on guessing unknown words. The present study therefore was an investigation of whether reading a Japanese summary that served as L1 background content knowledge affected Japanese L2 learners' ability to guess unknown words. The primary aim was to address the gap in the literature by building on the few available studies of how L1 background content knowledge impacts English as a Foreign Language (EFL) learners' ability to guess unknown words in written text.

Guessing Words from Context

Guessing from context is often used interchangeably with inferring from context. Researchers have expressed the idea in various ways, including "deriving word meaning from context," and "informed guessing from context" (Nation & Webb, 2011, p. 77). Schmitt (2000) stated that guessing an unknown word's meaning from context is known as inferencing from context (p. 152). Notwithstanding the different ways of referring to it, guessing from context is "the ability to derive a meaning for a word from context clues," excluding "unguided random guessing" (Nation & Webb, 2011, pp. 77-78). This definition of guessing from context was used in the present study.

Guessing from context is described as cumulative learning (Nagy, Herman, & Anderson, 1985; Nation, 2013). Nation (2013) noted that previous studies of guessing from context by L2 learners did not show impressive results partly because of the study designs and also because of "the effect of the cumulative nature of such learning involving only small gains per meeting for most words" (p. 354). Nagy et al. (1985) noted that the learning involved small increases in knowledge of a word and that learning from context is "a cumulative process where meaning and knowledge of form are gradually enriched and strengthened" (p. 355).

Nation (2013) pointed to the importance of taking into account realistic and desirable conditions for L2 readers to guess from context when investigating the proportion of unknown words that can be guessed. He stated that one such condition is that learners should already know a substantial number of words in a written text so that they can access the clues provided by those known words. The density of unknown words desirable for success-

ful guessing or adequate comprehension was recommended as one unknown word in 20 (95% density) (Liu & Nation, 1985) or one unknown word in 50 (98% density) (Hu & Nation, 2000; Nation, 2006). In other words, knowing 95%-98% of the vocabulary in a text is essential for unassisted comprehension to occur (Hu & Nation, 2000; Nation, 2006) and for guessing from the context (Hirsh & Nation, 1992).

Background Knowledge

The role of the reader's background content knowledge in reading comprehension is a topic that has received considerable attention from both L1 and L2 reading researchers. In L1 research, Wittrock, Marks, and Doctorow (1975), in a study with 484 elementary school students of English native speakers, investigated whether the meanings of undefined low-frequency words could be generated if the story context was familiar and meaningful to the participants. They found that those students who acquired familiarity with the story on their first reading generated the meanings of the low-frequency words inserted in the same text on the second trial. This led them to conclude that the learners were able to use their familiarity with the stories to produce the meanings of the low-frequency synonyms.

Other researchers investigated the interactive effects of prior knowledge and vocabulary difficulty on L1 students' recall of a reading text, finding that these two factors did not interact, or that there was no significant compensatory interaction between children's prior knowledge and vocabulary knowledge (Stahl & Jacobson, 1986; Stahl, Jacobson, Davis & Davis, 1989). Stahl, Jacobson, Davis, and Davis suggested that prior knowledge, or topic familiarity, affected participants' overall assignment of importance whereas vocabulary difficulty affected recall of both central and supporting information and the recall of the order of major events.

In L2 research, both general topic background knowledge and culturally oriented background knowledge play a facilitative role in L2 reading comprehension (e.g., Hayashi, 1999; Pulido, 2004). Grabe (2009) remarked that readers with considerably more background knowledge on a topic read a text differently and more efficiently with the proviso that the role of background knowledge in reading is not as simple as the past research had demonstrated. Schmitt, Jiang and Grabe (2011) also found that background knowledge had a facilitative effect on reading comprehension.

Guessing Words from Context and Background Knowledge

Engelbart and Theuerkauf (1999) provided a definition of context within vocabulary acquisition by taking into account a differentiation between verbal and non-verbal context. According to their definition, context comprises verbal and non-verbal contexts: Grammatical and semantic clues constitute verbal context; situative, descriptive, subject and global knowledge constitute nonverbal context. Background knowledge in the present study falls under the category of the subject context in the nonverbal context defined as "the information the learner already has in a given subject area" (p. 65).

Pulido (2004) pointed to the scarcity of research investigating the role of background knowledge in L2 incidental vocabulary acquisition despite empirical support for the positive effects of background knowledge on L2 text comprehension. Among the few existing studies, Qian (2004) found that there was a discrepancy between participants' views of their behavior when guessing unknown words and their actual

behavior. He found that young adult Korean and Chinese ESL learners relied more on the immediate semantic context and the forms of unknown words to obtain the meaning of the unknown words than on top-down strategies, which they thought they used most frequently.

On the other hand, Pulido (2007) referred to many studies observing L2 learners' strategies and knowledge sources used to infer word meanings during think-aloud protocols; the findings generally indicated that L2 learners, regardless of their proficiency level, tended to rely on background knowledge to guess word meanings (p. 68). Pulido also found that topic familiarity was conducive not only to guessing from context but also to retention.

When English learners read an English novel, they face both lexical and schematic challenges (Hammadou, 2000). While Laufer (1997) acknowledged the usefulness of L2 reader's background knowledge of the subject matter of the text, she cautioned that it could be an impediment to correct guessing and to reading comprehension when there is a discrepancy between reader schemata and the information in the text. Laufer and Sim (1985a, 1985b) pointed out that when learners interpret texts, they display a tendency to rely less on background knowledge and more on vocabulary knowledge.

Hayashi (1999) investigated the effect of extensive reading on developing reading strategies, such as skimming, scanning, using background knowledge and contextual guessing with 100 Japanese sophomores. She found that the use of contextual guessing increased, regardless of proficiency levels, after five months of extensive reading in which the participants were asked to read more than 100 pages per month. She also commented that having L1 background knowledge helped students comprehend the text with difficult vocabulary and grammar. Pointing to the significant difference the amount of L1 reading between the beginning level and intermediate students, she also concluded that reading extensively in the L1 might have provided the students with background knowledge that helped them comprehend L2 texts better.

Previous studies have indicated that background knowledge facilitates L2 learners' reading comprehension (e.g., Hayashi, 1999; Pulido, 2004); however, the role of background knowledge and the role that L1 background knowledge plays in guessing L2 vocabulary from context has been rarely investigated. This is an important gap because incidental learning through guessing from context is essential for vocabulary learning (Nation, 2001; Schmitt, 2000). The primary purpose of the present study therefore is to investigate L2 learners' ability to utilize L1 background knowledge when guessing the meaning of unknown words in reading texts.

Previous researchers investigating reading strategies have typically used short texts or genres such as newspaper articles, academic texts, and expository texts. Few researchers have used relatively long narrative texts. In the present study a relatively long narrative text was used to examine whether reading a Japanese summary that served as background content knowledge improved Japanese L2 learners' ability to guess unknown words.

The following research questions guide the present study:

1. What effects does readers' L1 background content knowledge have on guessing unknown and partially known words?

2. Do L2 learners differ in their guessed

unknown words if they read the L1 summary of L2 text as background content knowledge?

3. How do L2 learners use background content knowledge that was gained in their L1 for reading an English text with unknown vocabulary?

Method

Participants

Twenty-nine second-year students (13 female and 16 male students; mean age = 20 [SD = .78]) attending a four-year university in eastern Japan participated in this study. All of them had studied English for six years in junior and senior high school. All of the students volunteered to participate in the study. No standardized measure of their language proficiency was available; however, according to the participants' teacher, their English reading proficiency levels are intermediate. The participants were taking a 90-minute English class, twice a week at the time of the study.

The participants were randomly assigned to three reading conditions with or without the provision of an L1 summary: The first group (Pre-Reading Summary Group) read a summary written in Japanese before reading an English text. The second group (Pre- and During-Reading Summary Group) read the same L1 summary before reading and had it available while reading the English text. The third group (Comparison Group) just read the English text.

Research Design

The participants were randomly assigned to three reading conditions. The Pre-Reading Summary Group (n = 10) read the Japanese summary before reading the English text from *Animal Farm*; The Pre- and During-Reading Summary Group (*n* = 9) read the Japanese summary before and while reading the *Animal Farm* text: and the Comparison Group (n = 10) read the *Animal Farm* text without seeing the L1 summary.

Materials

An English text taken from Chapter 1 of the English novel *Animal Farm* (Orwell, 1945) was used in this study for three reasons. First, Pellicer-Sánchez and Schmitt (2010) demonstrated incidental vocabulary gains from authentic reading. Second, Pellicer-Sánchez and Schmitt also noted that novels that involve extended reading provide learners with an opportunity to be exposed to mid-frequency vocabulary. Third, it was important for the written text to contain enough low frequency words from which target words could be chosen. Real words were used as target words in order to ensure ecological validity (Nagy, Herman, & Anderson, 1985). *Animal Farm* was compared with another novel in a pilot study with six university students. It was determined that *Animal Farm* was more interesting and more comprehensible.

Nine low frequency words (e.g., *scullery* and *mare*) in the original text deemed too difficult for the participants were replaced with higher-frequency words (e.g., *kitchen* and *female horse*). One phrase (i.e., *in spite of the fact that his tushes had never been cut*) in the original text deemed too difficult for the participants was removed. The readability of the passage was then checked: The Flesch-Kincaid Grade level was 10; and Flesch-Kincaid Reading Ease score was 60. This score indicated that the text is easily understood by 13- to 15-year-old native English speakers. It is estimated that native speakers of this age-range have a vocabulary of around 11,000 to 13,000 word families based on a rough estimation that

native speakers' increase their vocabulary size by an average of 1,000 word families a year in their early life (Nation, 2001, p. 9). According to Hirsh and Nation's (1992) estimate, a vocabulary of around 5,000 words is necessary for L2 learners to read a novel written for teenagers who are native speakers of English (p. 60).

Based on Liu and Nation's (1985) and Nation's (2001) suggestions, 95% coverage or 98% coverage must be achieved if successful guessing from context is to take place. Nation (2006) estimated that "A vocabulary of 8,000 to 9,000 words is needed to read a novel, and even then, 1 word in 50 will be unfamiliar" and concluded that dealing with written text will require a 8,000-9,000 word-family vocabulary if 98% is considered to be the ideal coverage (pp. 71-79). Although the text contains quite a few proper nouns, such as *Manor Farm*, *Boxer*, and *Clover*, these proper nouns were treated as known words because readers can come to understand proper nouns as they read (Hirsh & Nation, 1992, p. 691).

Nation and Heatley's (2002) Range software using the British National Corpus (BNC) lists indicated that the 2,000-word level provided coverage of 90.05% of the text. Adding the 3,000-word level increased the coverage to 91.71%, and adding the 4000-word level further increased the coverage to 95.36%. The proper nouns in the text accounted for approximately 1.51% of the running words. There were 1,387 running words in the text. The time necessary for reading the text was found to be approximately 15-20 minutes in a pilot study.

A summary of the story was prepared in Japanese, the participants' L1, to be used for background information. The summary did not contain any of the target words in order to avoid pre-teaching them, but it contained the Japanese translations of other unfamiliar English words and idioms, such as *chew the cud*. To avoid using difficult *kanji* (Chinese characters) that might hinder the participants' comprehension of the Japanese translation, ten potentially difficult *kanji* that were identified in a pilot study were accompanied by easily readable phonetic transcriptions that were printed alongside the *kanji*. Three versions of the Japanese summary were piloted with three students and two adult learners of English so that they could choose the one in which the Japanese wording was the most appropriate and accessible. The time necessary for reading the Japanese summary was piloted with the same persons mentioned above and estimated to be approximately 4 minutes.

In order to address the time-on-task issue for Comparison Group, who did not read the Japanese summary, an English timed-reading text was chosen for them to read while the other groups read the Japanese summary.

Target Words

Thirty-three target words were chosen from the low frequency words in the text. Words that were similar to or regarded as Japanese loan words were excluded from the target word candidates. Waring and Takaki (2003) suggested that the guessing-from-context task should include at least 25 words to provide reliable results (p. 136). Using Nation and Heatley's (2002) Range software, four words were selected from the second 1,000 word frequency level of the BNC (i.e., BNC-2), six words from BNC-3, 18 words from BNC-4, BNC-5 and BNC-6; and five words from BNC-7, BNC-9, BNC-10, and BNC-16. The low frequency target words were selected based on the assumption that the participants were unlikely to have learned them. The four high frequency BNC-2 words were included as target words so as not to demotivate the

participants while completing the test and not to overwhelm them with an overly difficult test. The 33 target words were piloted with a sample of students of slightly lower and slightly higher proficiency. The mean score on the multiple-choice recognition test for the slightly higher-proficiency university students was 19 out of 33 (range = 12-25). The mean score on the same test for the slightly lower-proficiency university students was 10.95 (range = 2-19). These results indicated that there would likely be no floor or ceiling effects in the main study.

By considering previous research findings, 16 verbs, eight nouns, seven adjectives, and two adverbs were selected as target words. Liu and Nation (1985) found that adverbs and adjectives are more difficult to guess than nouns and verbs, and various studies have shown that nouns are the easiest to acquire, followed by verbs and adjectives, and then adverbs (Laufer, 1997). To control the role of frequency of occurrence and thereby make the role of background content knowledge more salient, most of the target words appeared only once, except for *foal*, *perch*, and *stout,* which appeared twice, and *comrade*, which appeared five times.

Measurement Instruments

A pretest was not administered in order to avoid pre-teaching the target words that would appear in their reading in the present study. To compensate for this, a prior-knowledge questionnaire was used to gather data on self-perceived prior-knowledge about the target words.

Word-form recognition test.

Two tests were administered immediately after the participants finished reading. One was a word-form recognition test modeled on Waring and Takaki (2003), which they used to "investigate the level of guessing" (p. 138). Upon collecting data on "learning from context" (p. 133), they recommended that tests should be sensitive enough to capture even a small amount of learning about new word forms and that a word-form recognition test allows participants to demonstrate such learning. The word-form recognition test in the present study contained the 33 target words plus an additional 21 distractors such as *toddler*, *spin*, and *farming*. The ratio of distractors to all words on the word-form recognition test (38%) was in line with Waring and Takaki's (2003) word-form recognition test (p. 140). The participants were asked to circle any words they thought they encountered while reading the text, or remembered seeing from the text that they read.

Bilingual multiple-choice recognition test.

The second test, administered immediately after the prior-knowledge questionnaire, was a bilingual multiple-choice recognition test modeled on Waring and Takaki (2003). This test was designed to assess correct inferencing. Nation (1990) suggested that the learners' first language could be used for correct answers and distractors to avoid a situation where students could not understand the meaning of the correct answers or the distracters written in English. The multiple-choice recognition test was, therefore, written in Japanese to avoid such a situation. An example item from the bilingual multiple-choice recognition test is shown below along with an English translation (Figure 1). Each target word was presented in isolation, and for each entry there was one correct answer and four distractors with an *I don't know* option added to reduce the effect of guessing (Waring & Takaki, 2003; Zhang, 2013).

Figure 1:
Example test item from the bilingual multiple-choice recognition test

Q 1. benevolent	1. 悪い	2. 太った	3. 気難しい	4. やさしい	5. わからない

benevolent	bad	fat	hard	kind	I do not know

Prior knowledge questionnaire and the reading task questionnaires.

The prior knowledge questionnaire and reading task questionnaires were administered immediately after the participants completed the reading. The prior knowledge questionnaire was a self-report of self-perceived word knowledge prior to the reading. The reading task questionnaires were intended to obtain data on how the participants felt about the level of difficulty in reading the English text, what they thought were the causes for the difficulty, whether reading the Japanese summary beforehand (Pre-Reading Summary Group) and before and during (Pre- and During-Reading Summary Group) was helpful, and how many English books they read a month in order to determine whether or not they read English books habitually. These questionnaires differed slightly, depending on the treatments the participants received. The final section of these questionnaires elicited personal comments about reading the English text.

Procedure

The participants were informed of the purpose of the study immediately after they finished the reading task questionnaires. Each participant' consent to participate was gathered using a written consent form.

All instructions were given in Japanese, as this was how the classes were normally conducted. The researcher was present to assist the teacher if necessary. The participants were not permitted to use dictionaries throughout the study. Administering the tests and questionnaires took one hour.

First, the Japanese summary was distributed to each participant in the Pre-Reading Summary Group and Pre- and During-Reading Summary Group while the timed reading material was distributed to the participants in Comparison Group. All participants were asked to read the text silently, which took about five minutes. The participants were also told not to write in the text and to turn over the text when they finished reading. They were also told that they would read a second English text, but they were not told if the text would be about the same topic.

Five minutes later, the teacher made sure that everyone had finished reading, asked the participants to stop reading, and collected the Japanese summary from the Pre-Reading Summary Group and the English text from Comparison Group. Only the Pre- and During-Reading Summary Group kept the Japanese summary so that they could refer to it if they wished while reading *Animal Farm*.

Next, the *Animal Farm* text was distributed

to all participants. The participants were asked to read it silently in 20 minutes or less, and to turn it over when they finished reading. They were also told not to write on the text and that there would be a test after the reading. After the participants finished reading, two immediate tests and two questionnaires were administered to all participants in the following order: word-form recognition test, prior word knowledge questionnaire, bilingual multiple-choice recognition test, and reading task questionnaire. Each time, the participants were reminded that the test scores would not influence their course evaluations and that adequate time would be provided for them to answer all items.

Marking

The correct answers on the word-form recognition test and the bilingual multiple-choice recognition test were awarded one point. The target words the participants knew and did not know were examined because it was assumed that each participant had different prior knowledge. The correct answers for each word on the bilingual multiple-choice recognition and the word-form recognition tests were calculated, and the results were compared to the scores on the prior knowledge ques-

Table 1

Summary of interpreting score combinations

Score combinations	Description	Interpretation
1-0-0, 1-1-0, 1-0-1, 1-1-1, 1-0-2, and 1-1-2	Correct answer in the multiple-choice recognition test. The participants did/did not recognize the word form. The participants had no/partial previous knowledge about the word.	Guessing unknown or partially unknown words because correct meanings were chosen in spite of no/partial word knowledge.
0-0-3 and 0-1-3	No correct answer in the multiple-choice recognition test. The participants perceived that they knew the word very well, but gave a wrong answer in multiple-choice test.	Unknown words because correct meanings were not chosen in spite of participants perceived full word knowledge.
0-1-0, 0-0-0, 0-0-1, 0-0-2, 0-1-1, and 0-1-2	No correct answer in the multiple-choice recognition test. The participants indicated that they did not know the word very well, and gave a wrong answer in the multiple-choice test.	Unknown words because correct meanings were not chosen and participants perceived they had no or little knowledge about the word.
1-0-3 and 1-1-3	Correct answer in the multiple-choice recognition test, but the participants did/did not recognize the word form. The participants indicated that they knew the word very well.	Known words because correct meanings were chosen and participants perceived full word knowledge.

Note. The order of score combinations from left to right as in 1-0-0 is as follows: bilingual multiple-choice recognition test (1), word-form recognition test (0) and prior word knowledge questionnaire (0) in the case of 1-0-0.

tionnaire (0 = Didn't know, 1 = Saw it before but did not know the meaning, 2 = Checked/learned it before but did not remember the meaning, 3 = Knew very well).

If a participant answered *stout* correctly on the bilingual multiple-choice recognition test and recognized the word on the word-form recognition test, and chose 0 for the same word on the prior knowledge questionnaire, the score combination was 1-1-0 for the target word. If a participant answered *wisdom* correctly on the bilingual multiple-choice recognition test, but did not recognize the word on the word-form recognition test, and chose 3 for the same word on the prior knowledge questionnaire, the score combination was 1-0-3 for the target word. On the other hand, if a participant answered *cheep* incorrectly on the bilingual multiple-choice recognition test and did not recognized the word on the word-form recognition test, and chose 3 for the same word on the prior knowledge questionnaire, the score combination was 0-0-3. Likewise, if a participant answered *conceal* incorrectly on the bilingual multiple-choice recognition test and recognized the word on the word-form recognition test, and chose 3 for the same word on the prior knowledge questionnaire, the score combination was 0-1-3. Neither of these cases was treated as known because a wrong answer was given in the bilingual multiple-choice recognition test. Table 1 summarizes how the score combinations were interpreted.

Results

One student was excluded from the study. This participant was regarded as an outlier because the number of correctly guessed words was 14 words, which produced a *z*-score of 2.94 (i.e., > 2.5). There were 28 remaining participants with 10, 9, and 9 participants in the Pre-Reading Summary Group, Pre- and During-Reading Summary Group, and Comparison Group, respectively.

Time data

Every participant read the Japanese summary (for Pre-Reading Summary Group and Pre- and During-Reading Summary

Table 2
Descriptive Statistics for Correct Answers in the Word-Recognition Test by Group (Max = 33)

	Pre-Reading Summary (n = 10)	Pre- and During-Reading Summary (n = 9)	Comparison (n = 9)
M	12.40	12.78	11.11
SE	2.17	1.96	1.87
95% CI	[7.49, 17.31]	[8.27, 17.29]	[6.79, 15.43]
SD	6.87	5.87	5.62
Skewness	.26	1.12	.12
SES	.69	.72	.72
Kurtosis	.02	1.41	.26
SEK	1.33	1.40	1.40

Note. Pre-Reading Summary Group = L1 Summary before reading the English text; Pre- and During-Reading Summary Group = L1 Summary before and while reading the English text; Comparison Group = the English text only.

Group) and the English text (for Comparison Group) within 6 minutes and the *Animal Farm* text within 20 minutes.

Word-form recognition data

The descriptive statistics for the correct answers given to the word-form recognition test by group are shown in Table 2. The means of Pre-Reading Summary Group and Pre- and During-Reading Summary Group were similar. Because of the small sample size, a Kruskal-Wallis test was conducted. The Kruskal-Wallis result showed that the number of the recognized words was not affected by reading a Japanese summary before or before and during reading the English text, H (2) = .25, p = .883. Among the target words, *benevolent* was recognized by five participants in the Pre-Reading Summary Group, two in the Pre- and During-Reading Summary Group, and one in the Comparison Group.

Data on guessing words based on the bilingual multiple-choice recognition data

In order to address the effects that participants' L1 background content knowledge had on their guessing unknown and partially known words (Research Question 1), I examined how many unknown words the participants in each group guessed successfully. Based on the aforementioned score combinations comprised of bilingual multiple-choice test results, word-form recognition results and prior-knowledge questionnaire, participants' known and

Table 3
Descriptive Statistics of Guessed Word for the Participants by Group (Max = 33)

	Pre-Reading Summary (n = 10)	Pre- and During-Reading Summary (n = 9)	Comparison (n = 9)
M	12.40	12.78	11.11
SE	2.17	1.96	1.87
95% CI	[7.49, 17.31]	[8.27, 17.29]	[6.79, 15.43]
SD	6.87	5.87	5.62
Skewness	.26	1.12	.12
SES	.69	.72	.72
Kurtosis	.02	1.41	.26
SEK	1.33	1.40	1.40

Note. Pre-Reading Summary Group = L1 Summary before reading the English text; Pre- and During-Reading Summary Group = L1 Summary before and while reading the English text; Comparison Group = the English text only

unknown words were estimated. Successfully guessed words were then estimated in the way mentioned in the previous section. Table 3 shows the descriptive statistics for guessed words. Means (with standard deviations in parentheses) of guessed words for Pre-Reading Summary Group, Pre- and During-Reading Summary Group and Comparison Group were 6.10 (2.38), 5.67 (3.08), and 4.11 (1.17), respectively. Means (with standard deviations in parentheses) for the ratio of guessed words to unknown words among the target words by group were 28% (0.10), 25% (0.12), 19% (.04), respectively. Performances of the best guessers by group were 10 words in Pre-Reading Summary Group, 11 in Pre- and During-Reading Summary Group, and seven in Comparison Group. Their unknown (and known) target words were 26 (7), 24 (9), and 26 (7) in number, respectively. Noteworthy is that the number of their known words (7, 9, 7) was not among the highest and was even lower than the mean of the participants, 10.18.

In order to address the difference in their guessed words with or without reading the L1 summary (Research Question 2), I examined if there was a systematic difference among the groups. Because of the small sample size, the Kruskal-Wallis test was conducted. The number of the guessed words was not affected by reading a Japanese summary before reading the English text, and before and during reading the English text, H (2) = 3.03, p = .22. Likewise, the ratio of the guessed words to unknown words was not affected by reading a Japanese summary before reading, and before and during reading the English text, H (2) = 3.27, p = .195.

Reading task questionnaire

In order to address how the participants used their L1 background content knowledge that was gained in Japanese for reading an English text with unknown vocabulary (Research Question 3), I analyzed their responses of questionnaires compared with those from Comparison Group. All participants responded to all the questionnaire items. The first questionnaire item (Q1), "Was the *Animal Farm* text you just read difficult to read?" was to examine participants' perceived ease or difficulty of reading the English text with 1 being *not difficult at all* and 5 being *very difficult*. The means (standard deviation) for the Pre-Reading Summary Group, Pre- and During-Reading Summary Group, and Comparison Group were 3.8 (1.07), 4 (0.87) and 4 (0.47), respectively; thus, most of the participants thought it was difficult. Q2 "For those who answered 1, 2 or 3, in the Q1, what do you think made it not so difficult for you?" was to elicit participants' reasons why they did not find the reading so difficult. Four out of ten participants in Pre-Reading Summary Group and two out of nine in Pre- and During-Reading Summary Group responded that it was manageable to read the text, and chose "I was able to guess the unknown words" and "I had Japanese background knowledge because I read the summary" as their reasons whereas one participant in the Comparison Group chose "I was able to guess the unknown words" and "I found the grammar/syntax accessible." On the other hand, those who found it difficult (who chose either 4 or 5 in Q1) chose "I did not know many of the words" and "I was not able to guess the unknown words" in Q3, irrespective of their group; thus, the reason for reading difficulty being unknown lexis. Two out of nine participants in the Comparison Group also chose "I did not have Japanese background knowledge," in addition to the lexical difficulties. Q4 asked if the participants found the Japanese summary conducive to reading the English story with 1 being *not helpful* and 4 being *very helpful* for the Pre-Reading Summary Group and Pre- and During-

Reading Summary Group. The means (standard deviation) were 3.40 (0.51) and 3.22 (0.44), respectively, indicating that it was helpful to read what they perceived as a difficult English story.

The most conspicuous differences among the different treatments were observed in the comments elicited in an open-response item where the participants wrote comments in Japanese about reading the *Animal Farm* text. I read through the data written in Japanese repeatedly and created categories (Brown, 2009). The comments fell into two broad categories, unknown words and comprehension. Most participants were concerned with the unknown words that they found were the stumbling block for comprehension. Some participants in the Pre-Reading Summary Group and Pre- and During-Reading Summary Group, however, wrote that they followed the main story line because they had read the L1 summary. Some participants in the Pre-Reading Summary Group wrote positive comments, such as "I was able to follow the story despite numerous unknown words because reading the summary was helpful to give me the gist of the story." Typical remarks included 'background knowledge', 'the Japanese summary', 'the context ', and 'unfamiliar vocabulary related to animals'. On the other hand, the Comparison Group's comments mainly concerned difficulty comprehending the story development, due to unknown words. Typical remarks included 'cannot concentrate', 'unable to grasp images of the main story line', 'unable to follow who did what'.

Some participants in the Pre- and During-Reading Summary Group were concerned with unknown words and phrases that were not in the L1 summary. Because the Japanese summary did not contain the target words, some participants noticed the gaps between the summary and the English text. They commented that the gaps occasionally contributed to confusion in their search for the equivalent meaning of the unknown English words in the summary although acknowledging that the summary in general helped them to read.

Specifically, the best guesser in the Pre-Reading Summary Group (10 guessed words) commented that meeting unknown words so often almost discouraged him to keep on reading. The best guesser in the Pre- and During-Reading Summary Group (11 guessed words) mentioned the difficulty in grasping animal characters and the detailed descriptions of their behaviors, which he said made the English text difficult to read. The best guesser in the Comparison Group (7 guessed words) commented that she was unable to comprehend the story from the beginning to the end due to a lack of knowledge of word meanings. These comments suggest that successful guessing did not seem to give them confidence.

Score combinations of the two tests and one questionnaire

Noteworthy was the occurrences of the score combinations comprised of the bilingual multiple-choice recognition test score (left), the word-form recognition test (center) and the prior word knowledge questionnaire (right) because all possible 16 score combinations occurred. It was not always the case that those participants who gave the correct answer to the target word on the bilingual multiple-choice recognition test were correct on the word-form recognition test because there were instances of the score combinations of 1-0-0, 1-0-1 and 1-0-2. The use of different tests of the same words provides different views of knowledge of those words. Even though the tests are differentially difficult, getting a word correct on a more difficult test does not mean that learners therefore

answer correctly on an easier test. This occurs because different test formats tap different kinds of knowledge and also trigger knowledge in different ways. The bilingual multiple-choice measure was the most difficult (receptive recall), the word recognition measure was next (receptive recognition) (see Waring and Takaki, 2003, for evidence of this), and the prior word knowledge questionnaire was the easiest.

Discussion

The present study was focused on the effect of L1 background content knowledge on guessing unknown words. There was no significant difference in guessing words with or without the L1 background knowledge; however, indirect evidence based on the participants' perception reflected in their comments in the reading-task questionnaire revealed that they felt that L1 background content knowledge slightly facilitated their comprehension of the story and might have helped them guess unknown words.

The Pre- and During-Reading Summary Group's large SD (3.08) for guessed words suggests that some participants benefitted from reading the English text with the presence of the Japanese summary. Based on the ratio of guessed words to unknown words, there appeared to be little difference in the degree of guessing unknown words between the Pre-Reading Summary Group and Pre- and During-Reading Summary Group. This indicated that referring back to the L1 background information while reading the English text did not give the participants in the Pre- and During-Reading Summary Group an advantage in guessing unknown words. A plausible explanation for this counterintuitive result concerns cognitive capacity. As Laufer (1997) stated, "Since the amount of information that can be cognitively manipulated at one point of time by controlled processing is limited, focusing on … words will take up some cognitive capacity that would otherwise be used for higher-level processing of the text" (p. 22).

The participants' comments on the open-ended items indicated that the participants' focus was mainly on words while their L1 background knowledge was a secondary focus. Laufer and Sim (1985a) investigated the "threshold of competence for L2 reading comprehension of academic texts," and concluded that vocabulary took precedence over knowledge of the subject matter for foreign language readers, regardless of their proficiency levels (p. 409). As Schmitt (2000) considers it important to train guessing because it is not automatically used by learners (p. 153), L2 readers might not be able to utilize their L1 background knowledge for guessing the meaning of unknown words unless they are trained to do so. It can be argued that if the participants had been trained to take advantage of their L1 background knowledge as a strategy for guessing, different results might have ensued.

In the present study, L1 background knowledge and the text content were similar for both the Pre-Reading Summary Group and Pre- and During-Reading Summary Group. Yet overriding contextual clues and suppressing them were observed in the form of "synform confusion" for many participants. Laufer (1997) calls errors caused by similar lexical forms "synforms" (p. 26). Apart from the lack of vocabulary, this might have been related to concepts. Some comments in the reading-task questionnaire referred to their unfamiliarity with novels and stories using personification written in English. Understanding what they perceived as many different animal names and detailed descriptions of animal behaviors, which they said were unfamiliar with, appeared to be especially problematic to quite a few for

these groups.

That said, those who appeared to have managed to follow the story's development were the participants in Pre-Reading Summary Group and Pre- and During-Reading Summary Group, many of whom remarked that reading the Japanese summary was helpful. Many participants in Comparison Group, on the other hand, commented that they were unable to comprehend the story because of too many unknown words. That the participants in Comparison Group did not guess as many words as the other groups and did not appear to comprehend the content as much as the others might be partially explained by their reading some unrelated English text while the other groups were reading the L1 summary of the English text.

The above findings echo the L1 research findings that background knowledge and vocabulary knowledge have different effects on reading comprehension (Stahl & Jacobson, 1986; Stahl, Jacobson, Davis & Davis, 1989). Stahl, Jacobson, Davis and Davis (1989) demonstrated that background knowledge seemed to have an effect on overall understanding of the passage whereas vocabulary difficulty had an effect on understanding particular parts of passage or particular clauses and sentences. If the vocabulary is difficult, the reader who has background knowledge tends to recall bits and pieces, not a coherent whole (Stahl, Jacobson, Davis & Davis, 1989). The results of the present study also indicated that background knowledge cannot compensate for the lack of vocabulary knowledge. In other words, that raising the L1 background knowledge did not result in a big increase in guessed words can be explained by their research findings, suggesting that background knowledge has a general effect on comprehension while guessing from context may require particular lexical knowledge. This is supported by Laufer's (1997) pithy statement that "words are the toys you need to play it right" when she was describing the indispensability of vocabulary knowledge to guessing and reading (p. 32).

It appeared, however, that knowing more target words beforehand did not necessarily work to learners' advantage in terms of guessing words. Even though some participants in the present study knew more target words, it did not necessarily translate into an increase in their guessed words. This supports the finding of Bensoussan and Laufer (1984) that "although better students may *know* more words, they do not guess differently or guess more than weaker ones" (p. 26). Gillis-Furutaka (2015) investigated what kinds of difficulties Japanese low-intermediate university and high school students faced when reading graded readers of their levels, and found that not only unknown words but also other factors emerged as causes for confusions and for impeding students' comprehension. She noted that even if the students understood every word in the text, they reported that they sometimes did not comprehend the content, due to other factors, such as "cultural differences" (p. 9). Likewise, Schmitt, Jiang and Grabe (2011) concluded that knowing vocabulary would be essential for greater comprehension, but other reading skills also played a role in facilitating comprehension, notably background knowledge.

A possible pedagogical implication would be raising learners' awareness about using their L1 background knowledge strategically to make most of the guessing strategy introduced by Liu and Nation (1985), that is to say, focusing on verbs and nouns by taking advantage of the context. Another suggestion would be the timing of providing background materials, if any. They could be provided before or before/while reading the authentic L2 text. It does not

matter whether the teacher allows them to keep them at hand while they read the English text. Again, training on utilizing more of their background knowledge may be desirable as the present study found that some guessing occurred when reading the English text that all participants perceived to be rather difficult, but their background knowledge was not activated enough to guess unknown words as in a case where "synform confusion" occurred.

Ideally, after reading, a follow-up activity would be necessary to confirm their correct guesses and rectify their wrong guesses. As the present study showed, participants tended to fall into a trap of "synform confusions" and applied their "preconceived notions" to words they did not know in the text just as Laufer's studies demonstrated (e.g., 1997). The Pre-Reading Summary Group's best guesser's comment suggested that guessing unknown words can be likened to walking in a dimly lit hall with a pair of dark sunglasses. Even if they successfully guessed, they would feel insecure and need reassurance afterwards. When students engaged in unassisted self-selected reading, Zimmerman (1997) found that students' perceptions about learning vocabulary appeared to "place more value on encountering words in natural environments than on memorizing them" (p.134). She also found that the same students who also received interactive word instruction as a treatment regarded such word instruction was more useful for learning words than self-selected reading.

Although further analysis is beyond the scope of the present study, all types of the score combinations occurred. We would expect a learner who got the largest previous knowledge measure to also get the multiple-choice and word recognition measures correct. We would also expect a learner who got the multiple-choice measure correct to also get the word recognition measure correct. In fact, every one of the 16 possible combinations occurred. It would be interesting to explore why particular learners got easier items incorrect and harder items correct. As Nation and Webb (2011) recommended as an elaborate way of displaying data from several tests, this variability in answering also might underline the importance of reporting results in detail and not just aggregating them.

Limitations of the Study

The results of the present study should be viewed by considering its limitations. The number of the participants in this study was quite small. This limits generalizability to other EFL learners at universities. A larger number of participants may have had more conclusive results.

An important limitation is not having used substitute or nonsense words, or foreign words that were unknown to the participants. Studies with substitute words or foreign words managed to control and measure unknown words precisely (e.g., Pellicer-Sánchez & Schmitt, 2010; Waring & Takaki, 2003). Although substitute words were not used in this study, the prior word knowledge questionnaire appeared to have compensated for this deficiency adequately. By the same token, controlling the unknown words and measuring the guessed words were not easy. Despite the comparable proficiency level, individual differences existed in their previously known vocabulary items among the target words. The self-reported prior word knowledge questionnaire and multiple-choice recognition test were the instruments used to measure which words were unknown to them because a pre-test was not administered in order to avoid preteaching. Selecting low-frequency words was an attempt to make sure the words

were unknown.

Another limitation is the effect of the I don't know option in the bilingual multiple- choice recognition test on test scores. Zhang (2013), suggested that the *I don't know* option not only reduces instances of guessing at the time of test-taking, but it also discourages partial knowledge. It is beyond the scope of the present study to know what effects that the I don't know option had on the participants while taking the test.

Finally, the results reported here describe intermediate Japanese learners of English in an EFL context. Learners with different native languages and higher proficiency may demonstrate different results.

Conclusion

The current study investigated what effect background knowledge had on guessing words. The material used was part of a narrative English text, *Animal Farm*. Administration of the treatments was done as a normal class activity in that the usual classroom teacher administered the treatments. The immediate word-recognition test, multiple-choice tests and two questionnaires were administered after reading. The target words were chosen from low frequency words from the English text. The participants in the study represented typical language learners in an EFL context in Japan and were intermediate learners.

Although the results showed no significant difference due to the presence of L1 background knowledge, but only indirect evidence of differences by treatment, it is premature to dismiss the importance of L1 background content knowledge relevant to the English content as useless. The participants across the different treatments did successfully guess a few unknown words after reading the English text that was perceived to be lexically difficult and conceptually unfamiliar. With or without L1 background knowledge, guessing occurred across the board. It seemed that having background knowledge helped moderately enhance participants' self-perceived comprehension of the text. Those with the L1 background knowledge followed the story more readily, compared to those who read the irrelevant English text. This can be explained by Stahl, Jacobson, Davis and Davis' (1989) study, suggesting that prior knowledge, or topic familiarity, affected participants' overall assignment of importance whereas the vocabulary difficulty affected recall of both central and supporting information and the recall of the major events.

As Schmitt and Carter (2000) see reading as a key component in most language learning programs, reading is integral to language learning. As they (ibid) remarked, intermediate learners would strive for the eventual transition from simplified readers for beginning students to authentic texts as they make progress to advanced levels. Learners will inevitably face a "lexical load" they have to cope with independently. Knowing about various strategies and equipping themselves with skills to deal with loads of new and unknown words is very useful.

Nagy, et al. (1985) stressed the importance of becoming "independent word learners" because "the number of words to be learned is too enormous to rely on word-by-word instruction" (p. 252). If EFL learners living in a country where English is not used on a daily basis continue to study English and wish to make progress from the intermediate to advanced levels, they must become "independent word learners" someday. Among reading skills, "guessing from context is an essential part of reading skill and essential prerequisite

for incidental vocabulary learning while reading" (Nation & Webb, 2011, p. 78). Background knowledge that is compatible with a reader's expectations and concepts contributes to successful guessing (Laufer, 1997). It is important to develop and train learners to guess because it is not automatically executed by learners (Schmitt, 2000). Aspirational "independent word learners" should be encouraged to train themselves to utilize L1 background knowledge as a strategy for L2 reading to help them guess from the context, which helps lead to comprehend the gist of the text.

Acknowledgments

I would like to thank Paul Nation for his advice and encouragement at various stages of this project and for reading and editing an earlier version of this paper. I would also like to thank David Beglar for carefully reading through, editing, and making constructive comments that have helped refine this paper. I also thank the two anonymous reviewers and the editor for their valuable comments that helped improve the clarity of this paper.

References

Bensoussan, M., & Laufer, B. (1984). Lexical guessing in context in EFL reading comprehension. *Journal of Research in Reading, 7*(1), 15-32.

Brown, J. D. (2009). Open-response items in questionnaires. In J. Heigham & R. A. Croker (Eds.), *Qualitative research in applied linguistics a practical introduction* (pp. 200-219). New York, NY: Palgrave Macmillan.

Clarke, D. F., & Nation, I. S. P. (1980). Guessing the meaning of words from context: strategy and techniques. *System, 8*(3), 211-220.

Engelbart, S. M., & Theuerkauf, B. (1999). Defining context within vocabulary acquisition. *Language Teaching Research, 3*(1), 57-69.

Gillis-Furutaka, A. (2015). Graded reader readability: Some overlooked aspects, *Journal of Extensive Reading, 3*, 1-19.

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice.* New York: Cambridge University Press.

Gu, Y., & Johnson, R. K. (1996). Vocabulary learning strategies and language learning outcomes. *Language Learning, 46*, 643-679.

Hammadou, J. A. (1991). Interrelationships among prior knowledge inference and language proficiency in foreign langugae reading. *Modern Language Journal, 75*, 27-38.

Hammadou, J. (2000). The impact of analogy and content knowledge on reading comprehension: What helps, what hurts. *Modern Language Journal, 84*(1), 38-50.

Hayashi, K. (1999). Reading strastegies and extensive reading in EFL classes. *RELC Journal, 30*(2), 114-132.

Hu, M., & Nation, I. S. P. (2000). Vocabulary density and reading comprehension. *Reading in a Foreign Language, 13*(1), 403-430.

Hirsh, D., & Nation, I. S. P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language, 8*(2), 689-696.

Joe, A. (1998). What effects do text-based tasks promoting generaiton

have on incidental vocabulary acquisition? *Applied Linguistics, 19*(3), 357-377.

Laufer, B. (1997). The lexical plight in second language reading: Words you don't know, words you think you know, and word you can't guess. In J. Coady & H. Huckin (Eds.), *Second Language Vocabulary Acquisition*. Cambridge: Cambridge University Press. 20-34.

Laufer, B., & Sim, D. D. (1985a). Measuring and explaining the threshold needed for English for Academic Purposes texts. *Foreign Language Annals, 18*, 405-413.

Laufer, B., & Sim, D. D. (1985b). Taking the easy way out: Nonuse and misuse of contextual clues in EFL reading comprehension. *English Teaching Forum, 23*(2), 7-10, 22.

Liu, N., & Nation, I. S. P. (1985) Factors affecting guessing vocabulary in context. *RELC Journal, 16*(1), 33-42.

Mondria, J-A., & Wit-De Boer, M. (1991). The effects of contextual richness on the guessability and the retention of words in a foreign language. *Applied Linguistics, 12*, 249-267.

Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly, 20*, 233-253.

Nation, I. S. P., & Heatley, A. (2002). Range: A program for the analysis of vocabulary in texts [Computer software]. Wellington, NZ: LALS, Victoria University of Wellington, New Zealand. Downloadable from http://www.victoria.ac.nz/lals/about/staff/paul-nation/nation.aspx.

Nation, I. S. P. (1990). *Teaching and learning vocabulary*. New York, NY: Heinle and Heinle.

Nation, I. S. P. (2001). Learning vocabulary in another language. Cambridge: Cambridge University Press.

Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge: Cambridge University Press.

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review 63*(1), 59-82.

Nation, I. S. P., & Webb, S. (2011). *Researching and analyzing vocabulary*. New York: Heinle Cengage Learning.

Orwell, G. (1945). *Animal Farm*. London: Secker and Warburg

Pellicer-Sánchez, A., & Schmitt, N. (2010). Incidental vocabulary acquisition from an authentic novel: do things fall apart? *Reading in a Foreign Language, 22*(1), 31-55.

Pitts, P., White, H., & Krashen, S. (1989). Acquiring second language vocabulary through reading: A replication of the Clockwork Orange Study using second language acquirers. *Reading in a Foreign Language, 5*(2), 271-275

Pulido, D. (2004). The Relationship between text comprehension and second language incidental vocabulary acquisition: a matter of topic familiarity? *Language Learning, 54*(3), 469-523.

Pulido, D. (2007). The effects of topic familiarity and passage sight vocabulary on L2 lexical inferencing and retention through reading. *Applied Linguistics, 23*(1), 66-86.

Qian, D. (2004). Second language lexical inferencing: Preferences, perceptions, and practices. In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language* (pp. 155-169). Philadelphia: Benjamins.

Saragi, T., Nation, I. S. P., & Meister, G. F. (1978). Vocabulary learning and reading, *System, 6*(2), 72-8.

Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.

Schmitt, N., & Carter, R. (2000). The lexical advantages of narrow reading for second language learners. *TESOL Journal, 9*(1), 4-9.

Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal, 95*(i), 26-43.

Stahl, S. A., & Jacobson, M. G. (1986). Vocabulary difficulty, prior knowledge, and text comprehension. *Journal of Reading Behavior, 18*(4), 309-323.

Stahl, S. A., Jacobson, M. G., Davis, C. E., & Davis, R. L. (1989). Prior knowledge and difficult vocabulary in the comprehension of unfamiliar text. *Reading Research Quarterly, 24*(1), 27-43.

Walters, J. (2006). Methods of teaching inferring meaning from context. *RELC Journal, 37*(2), 176-190.

Waring, R., & Takaki, M., (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language, 15*(2), 130-163.

Williams, R. (1985). Teaching vocabulary recognition strategies in ESP reading. *ESP Journal, 4*, 121-131.

Wittrock, M., Marks, C., & Doctorow, M. (1975). Reading as a generative process. *Journal of Educational Psychology, 67*, 484-489.

Wodinsky, M., & Nation, I. S. P. (1988). Learning from graded readers. *Reading in a Foreign Language, 5*(1), 155-161.

Zhang, X. (2013). The I don't know option in the vocabulary size test. *TESOL Quarterly, 47*(4), 790-811.

Zimmerman, C. (1997). Do reading and interactive vocabulary instruction make a difference? An empirical study. *TESOL Quarterly, 31*(1), 121-140.