The Role of Context-Based Mental Imagery in Intentional Vocabulary Learning Among Japanese EFL Learners

A Dissertation

Submitted to the University of Tsukuba

In Partial Fulfillment of the Requirements for

the Degree of

Doctor of Philosophy in Linguistics

Yusuke HASEGAWA

2014

Abstract

This study explores effective methods of teaching and learning vocabulary in English as a foreign language (EFL), focusing on (a) in what condition learners can take advantage of contextual information in vocabulary learning and (b) with what mental process learners read a context when their primary goal is to memorize the new words. Although some researchers believe that the combination of a translation and a minimal context, i.e., a single meaningful sentence, is the most economical way to learn the meanings of new words, it had been unclear how reading a minimal context would affect the learner's lexical knowledge. In examining the effect of context on vocabulary learning, this study focused on two specific factors. The first factor was learner proficiency, for prior studies indicated that only proficient learners can utilize contextual information effectively for vocabulary learning. In addition, this study focused on the second factor of *context imageability*, that is, the ease of evoking a mental image during reading a context. This factor seemed to be important in representing the semantic concept of the new words, translations, and contexts. Taking into account these factors, this study conducted six experiments, which were organized into three research projects, Studies 1 to 3.

Study 1 consisted of three experiments. As the first step, Experiment 1 explored whether context imageability affects vocabulary test scores when Japanese EFL learners memorized a word list consisting of 10 target words accompanied with their translations and contexts. According to the results of a lexical proficiency test, the high-intermediate learners were further divided into upper and lower level groups. As a result, although the upper group's scores were not different between the imageability conditions, the lower level learners scored higher in the more imageable condition than in the less imageable condition.

The lower group seemed more sensitive to the context type, presumably because they could not construct elaborate mental representations unless the contexts described a specific situation with concrete words. The results of Experiment 1 were fully replicated in Experiment 2, where participants were given *imagery instruction* that required them to imagine a situation described in each context. However, Experiments 1 and 2 did not provide any data concerning less skilled learners. Beginner level learners were recruited for Experiment 3, where more and less imageable context conditions were compared along with no-context conditions. Results showed that the more imageable condition was the most effective. However, when the participants were divided into upper and lower proficiency groups, the upper group was significantly influenced by the context imageability; further investigation was necessary with regard to how the proficiency factor interacts with the context imageability.

Study 2 examined two problems that emerged in Study 1. First, Study 1 assessed the participants' proficiency only with an English lexical proficiency test. Therefore, Experiment 4 considered not only learners' lexical proficiency but also their reading proficiency and learner types in terms of what information learners focus on the most during learning. However, the results suggested that the lexical proficiency measure was the most reliable in estimating the context imageability effect. Second, since participants in Experiments 1 and 2 were much more advanced than those in Experiment 3, it was necessary to examine the context imageability effect among participants at a medium proficiency level. Therefore, Experiment 5 compared the following three groups of learners: (a) high-intermediate, (b) low-intermediate, and (c) beginner groups. Results revealed the effect of context imageability was most significant in the high-intermediate group. It was suggested that even in a translation-based learning condition, contextual information is important for

high-intermediate EFL learners. In contrast, beginners and low-intermediate learners might depend on a given translation regardless of the given context. Considering the results from the upper group among the high-intermediate/advanced learners in Experiments 1 and 2, the context effect might also disappear among the advanced learners.

Given the results of Studies 1 and 2, Study 3 proposed some hypotheses as to how the context imageability factor affects cognitive processes involved in vocabulary learning. When highly imageable contexts are given to the learners, their mental process of context reading may include generation of mental imagery. In contrast, when contexts that are not imageable are provided for the learners, their mental process of context reading might be more analytic in nature; for example, their context reading might be characterized by more frequent use of syntactic and lexical knowledge. To examine the cognitive strategies during reading in more and less imageable contexts for the purpose of vocabulary learning, Experiment 6 used the *think-aloud method* that asked participants to report everything they came up with during learning. The think-aloud protocols and two kinds of supplementary data demonstrated that the cognitive strategies of context reading are different between more and less imageable conditions, fully supporting the hypotheses. Despite some limitations, the present research provides evidence suggesting that the cognitive strategies undertaken in intentional vocabulary learning using example sentences can be affected by the interaction between the learner's proficiency and the context imageability. According to the working model proposed in this thesis, teachers should not underestimate the role of contextual information in students' lexical development.

Acknowledgements

I owe a debt of gratitude to many people for their help and encouragement during my research. First and foremost, I am deeply indebted to Professor Yuji Ushiro, for his insightful comments, infinite patience, and invaluable encouragement. His meticulous suggestions were profusely valuable throughout the course of my research.

I would like to express my gratitude to Professor Hirosada Iwasaki, whose informative suggestions about the basis of the study were an enormous help to me, Professor Akira Kubota, whose careful comments helped me improve the quality of the thesis, and Professor Akiyo Hirai, whose thoughtful advice helped me to rethink the interpretation of the results. I am grateful as well to Professor Shuhei Kadota of Kwansei Gakuin University for acting as external examiner for my dissertation, whose perspicacious comments enabled me to improve the final draft.

I am pleased to acknowledge the support of the Japan Society for the Promotion of Science (JSPS) and the Society for Testing English Proficiency (STEP) Inc. for giving me a generous grant.

I would also like to express my appreciation to my senior associates, especially Dr. Yuko Hoshino, Dr. Chikako Nakagawa, Dr. Haruka Shimizu, Dr. Akari Kai, Dr. Shuichi Takaki, and Ms. Mayuko Kobayashi, for their technical support and helpful suggestions, and my junior associates, especially Mr. Akira Hamada and Ms. Yukino Kimura, for constructive opinions and warm encouragements. Many thanks are also due to my colleague and friend, Mr. Shingo Nahatame, whose constant cooperation and understanding of my work helped me over years. My family also deserves my gratitude for their hearty support. This work could never have been completed without help from them all.

Table of Contents

List of Tables

List of Figures

List of Appendices

Chapter 1

Introduction

1.1 Context of the Problem

Vocabulary in a foreign or second language (L2), as well as in the first language (L1), is acquired incrementally through various encounters in various contexts (Nation, 2013; Schmitt, 2010). However, compared to L1 words, which are acquired mainly through spoken language input, L2 words are learned also by deliberate memorization using translations. Many researchers in the fields of second language acquisition (SLA), psychology, and applied linguistics have studied the conditions in which learners can most effectively memorize new words in an L2 (see Section 2.2). The most typical kind of deliberate learning is to memorize word forms and meanings in association with one another using a list of L2 words with L1 translations. This kind of learning is called the *word association approach* (Jiang, 2000) and is known to be an effective method of learning unfamiliar words because it effectively maximizes how many words learners can retain after a limited time. In fact, researchers argue that the best way to learn L2 vocabulary is to first memorize the forms and meanings of new words deliberately and then learn more contextualized aspects of vocabulary such as collocation knowledge through extensive exposure in context (e.g., R. Ellis & Shintani, 2014). Through these combined learning methods, using translation and a number of contextual inputs, L2 learners can acquire word forms and meanings and develop intuition about many aspects of knowledge (see Nation, 2013, p. 49).

In contrast, learning vocabulary using context is called the *contextualized approach.* Examples of fully contextualized activities involve practicing the four language skills of reading, listening, speaking, and writing in authentic communication tasks (see Oxford & Scarcella, 1994). By reading an L2 book or listening to an L2 radio program, learners have

the opportunity to encounter numerous L2 words. Nowadays, there is a widespread view that vocabulary should be learned and tested through contextualized tasks (e.g., Qian, 2008; Read, 2000). However, vocabulary learning in context is a gradual process because a single context will supply information about only some aspects of the word's meaning (Schmitt, 2010). Actually, although most words have multiple meanings and referent concepts, only one of the many possible meanings of the word is often supported by the context. Learners should take advantage of both word association and contextualized approaches, because the first is more economical for L2 learners but the second is essential to enrich the learner's *mental lexicon*, i.e., the database in a learner's mind that contains information about orthographic, phonological, semantic, and many other aspects of words (see Kroll & de Groot, 1997, p. 180; see also Section 2.1.3). Thus, it seems obvious that the word association and contextualized approaches are in a mutually complementary relationship.

Researchers believe it is through context that various word meanings are conveyed and actualized (Nation, 2009). Considering this, even in the deliberate learning of word forms and their meanings using translations, context presentation seems to be necessary. For example, when a Japanese-speaking student is to learn a new word *clutch* with a translation such as *nigiru*, it might be helpful to read an example sentence such as, *The festival was extremely crowded, so Emily clutched her father's hand very tightly to keep from getting lost*. It is popular among Japanese learners of *English as a foreign language* (EFL) to learn English words using a list of new words paired with Japanese translations and example sentences. This kind of deliberate memorization of words seems to reflect aspects of both the word association and contextualized approaches, although the amount of contextual input is quite limited. According to Jiang (2000), the presentation of such contexts is helpful for learners to construct a link between L2 form and concept; contextualized vocabulary learning is an essential approach in acquiring the mental representation of target words (see Section 2.1.3).

When an L2 word is learned along with a context, learners may not only associate the L2 and L1 translation pair but also associate with them knowledge of how the given meaning can be applied to a specific situation.

Through this type of learning, learners might connect *word-based knowledge* (e.g., the spelling(s) and sound of the L2 word and a plausible translation in the L1) with *context-based knowledge* (e.g., the mental representation of the meaning of the given context and how the to-be-learned word is used in context). In this thesis, this connection constructed in the learner's mind is called the *word-context association* (see Section 2.5). The importance of context-based knowledge in this regard seems obvious, because L2 learners need to know how words are used in context. In fact, researchers often consider that knowledge of a new word is fully acquired only when a learner can use that word in an appropriate context. The importance of learner knowledge of contextual use is demonstrated most convincingly in vocabulary tests; examinees are sometimes asked to write a sentence or associate a word with a sentence (e.g., Qian, 2008; Read & Chapelle, 2001; Wesche & Paribakht, 1996). For example, in Wesche and Paribakht's (1996) *Vocabulary Knowledge Scale*, where learner development of lexical knowledge is reported in five stages, the deepest level knowledge is described as follows: *I can use this word in a sentence:* _____. As seen in this description, L2 learners are supposed to learn not only the meanings of lexical items but also the relationship between words and context. However, it is problematic that no prior research has sufficiently ascertained when and how learners build such a relationship in their minds.

1.2 Introduction to the Present Research

The present research explores how presentation of context enhances vocabulary learning. Specifically, this study deals with the problem of whether and how the word-context

association is constructed when L2 learners memorize a list of unfamiliar word forms with translations and example sentences.

According to prior research, it has not been obvious that the presentation of context is helpful for L2 learners. With respect to the effect of presenting an example sentence, Webb (2007a) published an influential research article entitled "Learning word pairs and glossed sentences: The effects of a single context on vocabulary knowledge." He compared two methods of learning a list of new words: vocabulary learning using (a) a list of word forms and translations (i.e., *decontextualized* learning) and (b) a list of word forms, translations, and example sentences (i.e., *contextualized* learning). Based on the argument that reading context has an important effect on learning, one might suppose that the contextualized condition has an advantage over the decontextualized one. However, Webb found no overall difference between these two conditions; as Folse (2004) notes, context presentation is not as effective as people have believed. Thus, there is a gap between the assumption that context presentation is helpful for the deliberate learning of word forms and their meanings and the empirical finding that context presentation has no direct effect on vocabulary learning. To tackle this problem, the present study, launched in 2010, carefully designed experiments to examine context effects during vocabulary learning. Consequently, it was discovered that L2 learners are more sensitive to context information than Webb and other researchers had expected and that learners can take advantage of context if some conditions are satisfied.

The current research has two main focuses. The first is how teachers and learners can optimize the effect of context reading for vocabulary learning. Past studies have introduced only a limited amount of evidence showing in what condition L2 learners can take advantage of contextual input during vocabulary acquisition. Second, this study explores the nature of learner comprehension of context while their intention is oriented toward memorizing lexical items. To explain these matters, it was important to consider the mental connection between

lexical items and context information. This research hypothesized that, after learners encounter lexical items in context, the contextual information is retained in relation to the target lexical items. Some learners retain context by drawing a picture in their minds of the situation described in the context.

For example, suppose a learner reads a context such as *Mrs. Smith was a very* <u>*pious*</u> *woman who attended church services regularly*, and that learner is not familiar with the word *pious* (this material was used in the author's prior study, Hasegawa, 2010). Learners might retain *pious* in relation to an image aroused by the context concerning the described person, *Mrs. Smith*. If this is the case, then learners can recall the meaning of the target word *pious* in relation to the context and vice versa. In addition, if a learner retains the imagery from context in her or his mind, then this learner will perform better in a certain type of contextualized vocabulary test. For example, such imagery will work effectively in a contextualized test such as the following, where the multiple choices are (a) *adjoining*, (b) *bogus*, (c) *pious*, and (d) *sporadic*: "Choose the best word for the blank in the sentence. *Mrs. Smith was a very* ( ) *woman who attended church services regularly*." Activating the mental association between *pious* and the image of the described person *Mrs. Smith*, a learner may recall that the target item's meaning is related to a personality that best matches the described situation. Learner performance in such a contextualized test has never been investigated in light of mental imagery representation from context.

To examine the function of context reading during vocabulary learning, and to gain pedagogical awareness as to how teachers and learners can make contextualized vocabulary learning more effective, this study conducted six experiments (referred to as Experiments 1 to 6) and four pilot studies (see Figure 1.1). In this thesis, these six experiments are organized into three research projects (Studies 1 to 3), as follows. Study 1 consists of three experiments that aimed to examine the effects of mental imagery from context on intentional vocabulary

learning. Through two experiments, Study 2 further examined the relationship between learners' L2 lexical proficiency and the effect of the image-evoking value of context (i.e., context imageability effect) on vocabulary test scores after contextualized learning. Based on the results of these studies, Study 3 proposed some hypotheses as to how the context imageability factor affects cognitive processes involved in vocabulary learning. All six experiments conducted for the three studies were closely interrelated; this thesis includes general discussion on the relationship between the results. In addition, this study was designed to provide pedagogical implications about whether and how teachers should use context in a classroom to teach L2 vocabulary effectively.

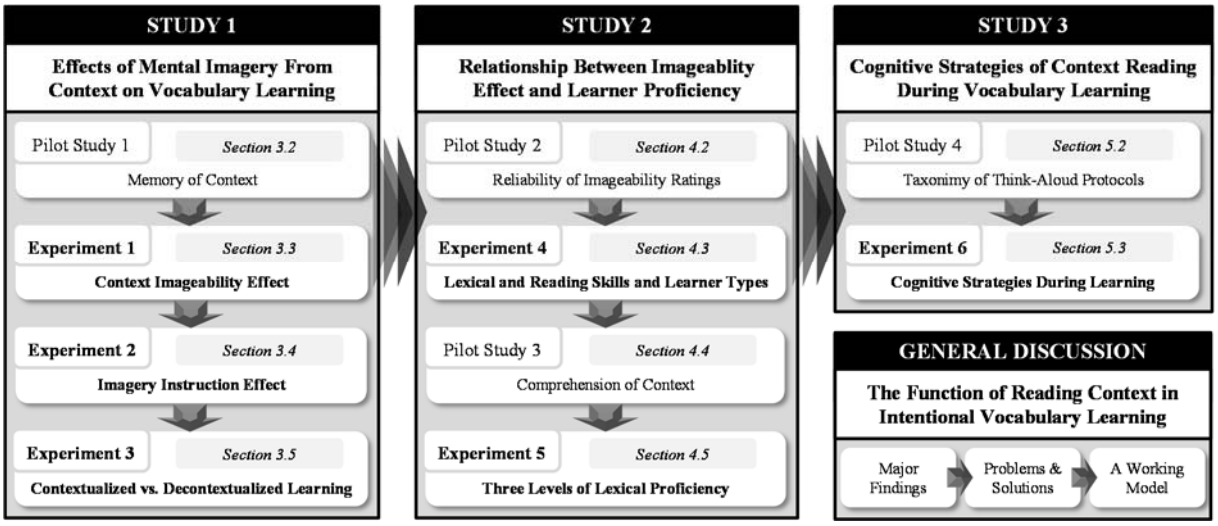

Figure 1.1. The overall scheme of the present research.

1.3 Organization of the Present Thesis

This thesis consists of seven chapters. Having introduced the context of the problem in this chapter (Chapter 1), Chapter 2 reviews the related literature on (a) vocabulary development in L1 and L2, (b) intentional and incidental vocabulary learning, (c) the process

of context comprehension, and (d) generation of mental imagery. This review is followed by a summary of the position of the present study in the research field of L2 vocabulary learning.

Chapters 3 to 5 report the purposes, methods, and results of the experimental studies. The organization of the main and pilot studies is displayed in Figure 1.1. Chapter 3 describes Pilot Study 1 in the order Experiment 1, Experiment 2, and Experiment 3. In a similar manner, Chapter 4 reports Pilot Study 2, which comprises Experiment 4, Pilot Study 3, and Experiment 5 in this order. A general design of the experiments in Studies 1 and 2 can be regarded as the *learning-testing paradigm*. Based on Nation and Webb's generalized diagram of the design of an experimental study (2011, p. 19), Figure 1.2 will be helpful for understanding the general design of Experiments 1 to 5. In each experiment, participants learned a list of unfamiliar words and then took a test on the target words. At least two learning conditions were compared in each experiment. Conducting a pretest was optional: If target words were pseudowords, or nonwords with disguised word forms, it was unnecessary to measure participants' prior knowledge of the target words. In addition, a delayed test of target words was administered when necessary. Chapter 5 explains Pilot Study 4 and Experiment 6, in which participants were asked to learn a list of unfamiliar words and report everything they thought and felt. Although Experiment 6 had a learning phase and a testing phase, as shown in Figure 1.2, the experimenter's interest was in participants' cognitive processes during learning, rather than posttest scores.

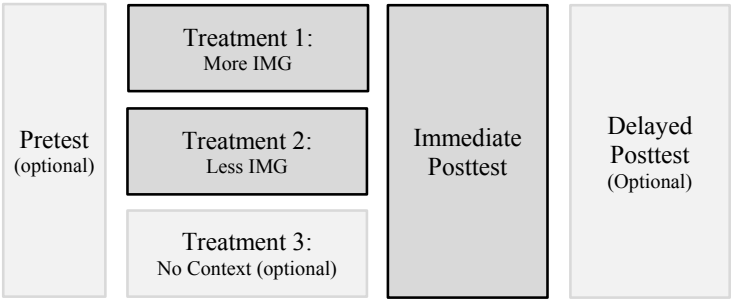

Figure 1.2. The general design of Experiments 1 to 5. IMG refers to imageability of context, which is introduced in detail in Chapter 2.

Based on the results from all experiments, Chapter 6 summarizes the main findings and discusses the theoretical and practical issues. In particular, this chapter focuses on how learners construct and utilize the imagery representation aroused through reading context and associate it with target lexical items. The discussion includes the introduction of a working model that explains the relationship between factors affecting contextualized vocabulary learning. The concluding remarks are made in Chapter 7. This chapter also clarifies the importance of the current research and discusses limitations. Finally, it outlines areas for future research and the pedagogical implications for intentional learning of vocabulary in context. The experimental materials and collected data from the experiments are presented in the Appendices. Information provided in the Appendices should help to concretize and clarify the discussion in each chapter.

Chapter 2

Literature Review

The present study is based on research findings and suggestions from both vocabulary learning and context reading studies. Therefore, this chapter covers both research fields as follows: review of literature on vocabulary learning, covering topics concerning (a) vocabulary development in L1 and L2 (Section 2.1) and (b) intentional and incidental vocabulary learning (Section 2.2), and literature on reading comprehension, covering topics concerning (c) the process of context comprehension (Section 2.3) and (d) generation of mental imagery (Section 2.4). This review is followed by a summary of the position of the present study in the research field of L2 vocabulary learning (Section 2.5).

2.1 Vocabulary Development in L1 and L2

2.1.1 Vocabulary Acquisition in L1

To some extent, research on vocabulary learning in L2 has been developed based on research findings and methodology used in studies of L1 vocabulary acquisition. For example, Nagy, Herman, and Anderson (1985) suggested that children may learn as many as 3,000 words per year between grades three and twelve. Children in elementary schools can acquire words through a number of encounters at a rapid rate, whereas only a small percentage is generally acknowledged to be learned through teachers' direct instruction. Therefore, in L1 vocabulary acquisition, children learn new vocabulary incidentally from listening and reading. There have been numerous investigations to support the belief about reading and vocabulary learning as an incidental outcome of reading (Swanborn & de Glopper, 1999, 2002). These studies found evidence showing that children could learn vocabulary indirectly while reading the words in context.

A number of studies on L1 vocabulary acquisition have shown that elementary school students can acquire thousands of lexical items through reading (see Krashen, 1989). The basic and typical framework of these experiments was as follows: Students were first given a text with some unfamiliar words to read but were not told to learn the new words. After reading the text, they were tested on the previously unfamiliar words without preannouncement of testing. The result of such studies repeatedly showed that readers make small but significant gains in vocabulary knowledge, from even one exposure to a new word.

Some researchers expect a similar pattern of vocabulary gain in L2 learning; one of the most influential examples is Krashen (1989). His claim, which is widely known as the *monitor model*, includes the *input hypothesis*: Comprehensible input may result in L2 acquisition. That is, a learner can acquire the target language through a comprehension process of the message conveyed in the L2 rather than the process concerning linguistic form. By extending L1 research findings to L2 learning, he suggested that an understanding of written context would result in good vocabulary acquisition. However, the evidence for L2 learners gaining vocabulary while reading is not as well established as it is for L1 learners. As stated in the next section, the situation for L2 (especially EFL) learners is not as easy as for L1 learners in terms of the amount of linguistic input received and other factors.

2.1.2 Vocabulary Learning in L2

The previous section mentioned that there is a difference between L1 acquisition and L2 or EFL learning when it comes to vocabulary. In the present study, the term EFL refers to a situation in which learners do not have to use English in everyday life and typically learn English only in language classrooms; in other words, English is learned in *educational settings* rather than *natural settings* (R. Ellis, 2008). In general, EFL learners are likely to be influenced by L1 transfer, intelligence, and motivation (Shirahata, Tomita, Muranoi, &

Wakabayashi, 2009). However, the most obvious difference is that the amount of linguistic input in the target language is smaller in the EFL context than the L1 context. Because the probability of gaining a significant amount of knowledge about a word through a single encounter of it in context is quite limited (e.g., Nagy et al., 1985), frequency of encounter is the most influential determinant of successful learning. Compared to the amount of L1 input, the total frequency of L2 input is considerably small in the EFL situation; it is unrealistic to suppose that learners can acquire vocabulary without any effort or support for learning.

Therefore, Nation (2013) suggests that the most effective method of L2 vocabulary learning is a combination of deliberate learning such as memorizing a word list using translations into the L1 and message-focused activities where learners read, listen to, and use new words in a communicative context. According to Nation, deliberate learning using translations seems to be more necessary for high-frequency words and technical terms. It is important for high-frequency words because learners can hardly understand the context unless they are familiar with most frequent words. Any type of contextualized learning requires learners to have knowledge of high-frequency words to some extent. For technical terms that are much less frequent, deliberate learning is highly likely to be helpful because students should learn them within a limited number of encounters. There are not enough chances to encounter these words in novels, newspapers, or other usual reading materials, whereas knowledge of these terms is inevitable for understanding the overall meaning of what the learner is going to read or listen to.

However, it is worth noting that Nation's definition of "high-frequency" words covers 2,000 *word families*; in Schmitt and Schmitt (2014), this range includes 3,000 word families, which roughly correspond to 5,000 *lemmas* (Mochizuki, Aizawa, & Tono, 2003). A word family differs from a lemma in how individual words are counted: For example, the word family for *access* is composed of its inflected and derived entries such as *accessed*, *accesses*,

accessing, *accessible*, *inaccessible*, *accessibility*, and *inaccessibility*, whereas the lemma for the verb *access* includes only *accessed*, *accesses*, and *accessing*, regarding its derived forms such as *accessible* and *accessibility* as different lemmas (Nation & Webb, 2011). Considering that the Course of Study in Japan currently requests only 3,000 lemmas to be taught in public education (Ministry of Education, Culture, Sports, Science and Technology, 2008, 2009), acquisition of all the high-frequency words in Nation's view might be challenging for Japanese EFL learners. Furthermore, the 3,000-word-families baseline might be unsatisfactory in terms of text coverage. According to Nation (2006) and Laufer and Ravenhorst-Kalovski (2010), 4,000 to 5,000 word families are necessary for 95% coverage of normal reading materials and 8,000 word families for 98% (see also Hu & Nation, 2000). Therefore, deliberate learning using translation is helpful for most EFL learners because their lexical knowledge seems necessary to be extended to achieve such a great vocabulary size.

2.1.3 Bilingual Mental Lexicon

Another significant difference between L1 and L2 vocabulary learning is that L2 learners have prior knowledge of the L2 lexical system; this is why Nation and Webb (2011) suggest that translation is the fastest way of presenting meanings of new L2 words. Although there is a widespread view that it is undesirable for teachers to use the L1 to explain the meaning of L2 words, there is in fact no evidence to support this claim (Laufer & Nation, 2012). Van Hell and de Groot (1998) proposed a framework called *Distributed Conceptual Feature Model* to explain how overlap of conceptual features between L1 and L2 words works in the word recognition process (see also de Groot, 1992; Kroll & de Groot, 1997). Figure 2.1 provides an overview of this model. An L2 word (e.g., *water*) has a number of conceptual features, and its L1 equivalent (e.g., *mizu* in Japanese) shares some of these features. They proposed that concrete nouns, whose referents are easier to imagine than

abstract nouns', likely share more conceptual features with their L2/L1 equivalent than abstract words. As Figure 2.1 shows, there are some gaps between L1 and L2 words; some features of the L2 word are not shared by its translation. To fill the conceptual gap and fully understand the true meaning or nuance of the L2 word, learners should encounter the word in a variety of contexts and incrementally learn it.

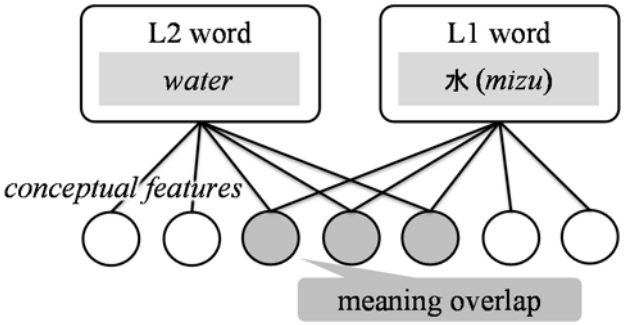

Figure 2.1. An example for Distributed Conceptual Feature Model adapted with some modifications from Kroll and de Groot (1997, p. 188).

For this reason, although L1 translation is a prompt source of learning, it is dangerous to learn L2 words without any contextual input. Furthermore, the importance of learning vocabulary in context can be emphasized in terms of the "fossilization" phenomenon (Jiang, 2000). In Jiang's (2000) theory, L2 learners' mental lexicon is modeled as in Figures 2.2 and 2.3. Figure 2.2 illustrates the lexical representation and processing at the final stage of lexical development. At this ideal stage, learners can understand and use L2 words without any intermediation of the L1; in Figure 2.2 (a), the content of their lexical representation is filled with the L2 rather than L1 entity. In addition, Figure 2.2 (b) illustrates that L2 word representation is directly linked to conceptual representation. For example, learners at this final stage of lexical development can form a mental representation of *cat*, *water*, or *table* when they read or listen to these words. The models illustrated as Figures 2.2 (b) and 2.3 (b)

are directly based on Kroll and Stewart's (1994) *Revised Hierarchical Model* (see also Kroll & de Groot, 1997). This is one of the best-known models of mental lexicon development related to a number of word recognition experiments (e.g., Kroll & Stewart, 1994; Kroll, van Hell, Tokowicz, & Green, 2010; Sunderman & Kroll, 2006). It is worth mentioning that Revised Hierarchical Model was originally designed to account for out-of-context translation performance. For example, if pictures are presented as a kind of semantic context, participants' performance on translation tasks becomes similar between L1-to-L2 and L2-to-L1 directions (see Kroll & de Groot, 1997, pp. 182–183).

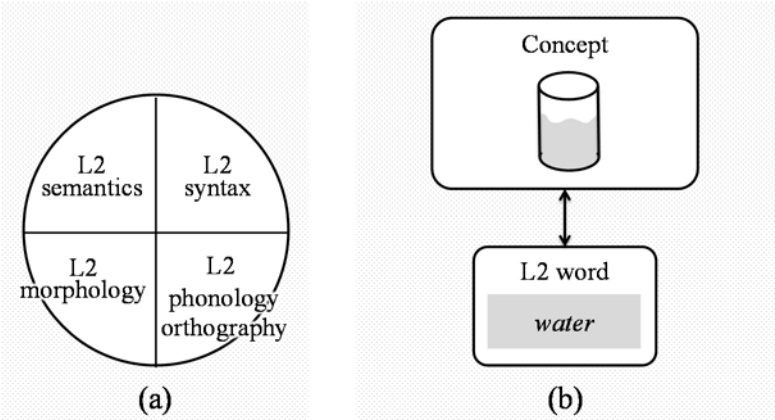

Figure 2.2. Lexical representation (a) and processing (b) at the final stage of lexical development in L2 (adapted with some modifications from Jiang, 2000, p. 53).

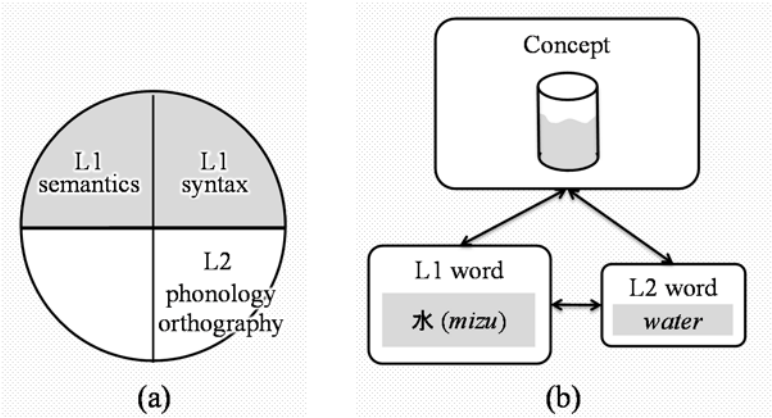

Figure 2.3. Lexical representation (a) and processing (b) at the L1 lemma mediation stage of lexical development in L2 (adapted with some modifications from Jiang, 2000, p. 53).

In principle, all learners may reach this level if there is a sufficient amount of contextualized input that they successfully process, according to Jiang (2000). However, unfortunately, the lexical competence in an L2 learner may fossilize, or stop developing, at a former stage as illustrated in Figure 2.3. At this stage, called the *L1 lemma mediation* stage, the L2 word meaning (i.e., semantic concept) is reached through the mediation of L1 translations. For example, the English word *water* can be understood only through the Japanese translation *mizu*, if a Japanese learner's lexical competence of this word is at the L1 lemma mediation stage. According to Jiang's account, a major cause of lexical fossilization is the lack of contextualized exposure.

Therefore, vocabulary learning in context is an important approach for the development of lexical competence and knowledge. From a pedagogical viewpoint, contextualized vocabulary learning is effective in that learners may encounter numerous words by reading and listening to passages. In this way, learners develop their fluency by encountering familiar words and extend their vocabulary size by encountering unfamiliar words (e.g., Nation, 2009). In summary, the learning mode that the present study focuses on (i.e., deliberate learning of a list of L2 word forms using L1 translations and example sentences) is assumed to be effective in reinforcement of the L1-L2 lexical connection. Referring to a review by Schmitt (2008), R. Ellis and Shintani (2014) also present the idea that the form-meaning link should be learned first through an explicit teaching approach, and the more contextualized aspects of vocabulary such as collocation knowledge are best learned implicitly through extensive exposure to the use of words in context (p. 107). However, the mechanism by which learners' cognitive process of reading context affects vocabulary learning is not fully applied or explained in any theoretical models; the present study explores how the cognitive process of reading context affects the development of the mental lexicon.

2.2 Intentional and Incidental Vocabulary Learning

2.2.1 Intentional Learning Tasks

There are a number of types or variations of vocabulary learning, so researchers have established a number of classifications to describe the nature of different types of vocabulary learning. This section briefly reviews the following three sets of features: (a) *intentional* vs. *incidental*, (b) *contextualized* vs. *decontextualized*, and (c) *translation-based* vs. *guessing-based* learning (see Figure 2.4). This section also defines the types of vocabulary learning that this study will focus on. Previously, the general term deliberate learning has been used in this paper to represent a situation where a learner uses a list of new words paired with their translations and memorizes them with some effort. This is the most typical example of intentional learning. This section further defines the learning mode that this study focuses on in terms of the three sets of features mentioned above.

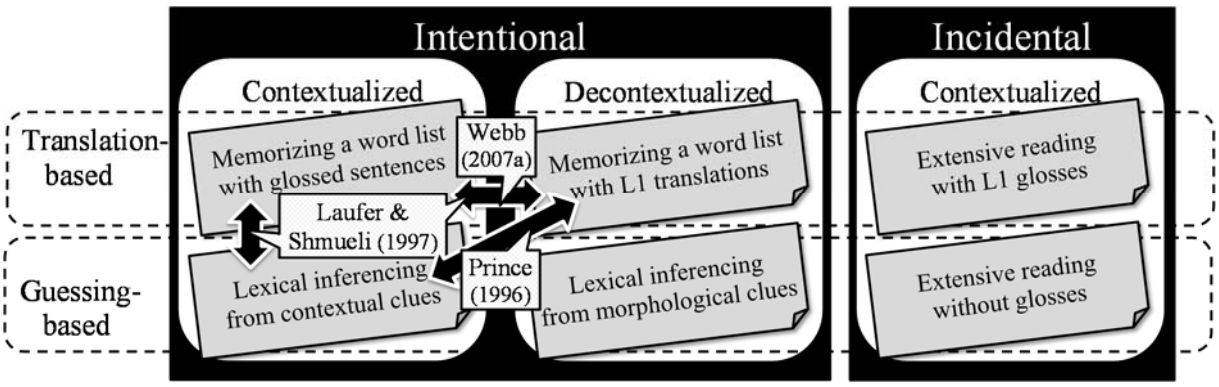

Figure 2.4. Classification of vocabulary learning. A typical task for each category is also presented. Black arrows show comparisons made in three past studies (Laufer & Shmueli, 1997; Prince, 1996; Webb, 2007a); the present study followed Webb's (2007a) comparison.

2.2.1.1 Classification of Intentional Learning Tasks

Researchers have distinguished intentional from incidental learning of vocabulary since the 1980s at the latest (Krashen, 1989; Laufer & Hulstijn, 2001). In spite of some arguments against the validity of this contrast (Bruton, Lopez, & Mesa, 2011), this classification has been approved by most researchers attempting to make a rough distinction between deliberate learning of lexical items (intentional learning) and skill-based or communicative activities, in which learners gain lexical knowledge as a byproduct of a main task (incidental learning). The present study focuses on the first mode of learning. According to Hulstijn (2005) and Schmitt (2010), the most typical intentional learning situation can be created when the learner's attention is placed on the memorization of a list of words and the learner is aware that there will be a test on the material learned; the present study realized this situation. Further discussion and summary in terms of incidental learning is presented in Section 2.2.2.

The second viewpoint from which the types of vocabulary learning can be differentiated is whether target words are presented in or out of context. Representative articles comparing contextualized and decontextualized learning within the intentional learning paradigm include Laufer and Shmueli (1997), Prince (1996), and Webb (2007a). Laufer and Shmueli, for example, compared intentional vocabulary learning in both sentential contexts and the context of a passage from a course textbook with learning under a no-context condition. Participants' scores on a delayed test dropped significantly under the condition that new words were presented in a full paragraph context. It was suggested that contextualized learning is more effective in a sentential context, while the decontextualized approach is also an efficient means of learning.

The third difference in classification can be observed in terms of how learners determine the meaning of target words. In particular, researchers have been interested in how knowledge of word meanings acquired from contextual information is different from that

acquired from translations. In an intentional learning study, Prince (1996) compared two conditions: a contextualized condition, where learners were presented with a list of sentential contexts that included unfamiliar words, and a decontextualized condition, where they were given a list of unfamiliar words paired with translations. Generally, the decontextualized method of learning seemed equally effective regardless of the learner group. In contrast, the results for contextualized learning were affected by both learner proficiency and test type as follows: Guessing-based contextualized learning was ineffective for learners with lower L2 proficiency; however, contextualized learning brought upper learners' scores on a contextualized test to a higher level, closer to those achieved in decontextualized learning.

To summarize, in terms of how many words learners can retain after a limited time, the most economical method of learning the connection of new word forms and their meanings is the intentional and translation-based approach. In addition, assuming that contextual information is helpful for learners to construct a link between word forms and their concept (see Section 1.1), presentation of example sentences seems to be necessary. Thus, the present study focuses on a learning mode that is (a) intentional, (b) contextualized, and (c) translation-based. However, it has been unclear how the context presentation affects vocabulary learning. As introduced in Section 1.1, the actual effect of context presentation was found to be quite small. In fact, when Webb (2007a) compared vocabulary learning using (a) a list of word forms and translations and (b) a list of word forms, translations, and example sentences, there was no clear difference between these two conditions. The next section discusses why context presentation was not as effective as expected by pointing out several problems amongst previous studies.

2.2.1.2 Problems in Previous Research

In fact, there have been many studies suggesting that contextualized learning is an ineffective or inefficient way of learning (see Folse, 2004). In particular, researchers such as Griffin (1992) and Prince (1996) have found that it is likely to be unsuccessful when the learner does not have enough proficiency in the target language. Interestingly, the "lower proficiency" participants in Prince's and other researchers' studies were not beginners; their typical proficiency level can be characterized as low- or high-intermediate, which suggests that most of them were able to comprehend the context but were not very fluent at processing that contextual information. On this point, recent studies have suggested that lower proficiency learners are incapable of associating contextual information with vocabulary knowledge (e.g., Nassaji, 2006). Therefore, the process of drawing information from context and associating it with vocabulary knowledge can be assumed to be quite difficult and to require considerable proficiency.

These results are best understood in terms of *aptitude-treatment interaction* (ATI; Cronbach & Snow, 1977), which refers to a situation where a particular instructional condition is more or less effective depending on the learner's abilities or characteristics (see also Vatz, Tare, Jackson, & Doughty, 2013). However, as Figure 2.4 shows, Prince's (1996) two learning conditions involved two different comparisons: not only decontextualized vs. contextualized but also guessing-based vs. translation-based learning. In other words, it was impossible to determine whether the ATI (i.e., the learning condition × proficiency group interaction) was caused by context presentation or guessing probability.

Unfortunately, there have been no clear evidence showing whether the effect of context presentation interacts with learner proficiency when the learning condition is strictly controlled to compare (a) intentional and translation-based learning in context and (b) intentional and translation-based learning apart from contexts. However, Webb (2007a)

compared contextualized and decontextualized learning groups, presenting both groups with translations: Participants in the contextualized group used a list of glossed contexts consisting of example sentences and translations of the target words. Over 10 tests, he examined the difference between the translation-only and translation-and-context conditions. The results of his statistical analysis showed no significant overall difference, and he concluded that a single glossed sentence as context may have little effect on vocabulary knowledge. However, one might wonder whether the absence of proficiency as a factor might have affected the results of his experiment; an effect of context presentation might have appeared if Webb's participants had been divided into proficiency groups.

Another momentous difference between the experiments by Prince (1996) and Webb (2007a) was the types of tests implemented. As mentioned earlier, Prince conducted both contextualized and decontextualized tests, both specialized to assess learner knowledge of form-meaning mapping. More specifically, the contextualized test required participants to recall a target word, e.g., *crush*, to fill in a blank in a sentence such as, "Wine is made by _____ grapes and letting the juice ferment" (Prince, 1996, p. 492). This was a new context for both the decontextualized (translation) and contextualized groups, as the latter group learned the target word *crush* in a different sentence: "Why are those eggs all broken?" "Well, someone put the eggbox on the chair, and I didn't see it, and I sat on it and <u>crushed</u> it." In contrast, using 10 tests, Webb measured five aspects of vocabulary knowledge (orthography, meaning and form, paradigmatic association, syntagmatic association, and grammar); however, only the two grammar tests were contextualized, with students being asked to write or recognize a grammatically correct sentence including a target word.

The contextualized recall format is useful for assessing a learner's incremental knowledge gain. According to Sonbul and Schmitt (2010), who adopted the contextualized recall test together with some decontextualized tests, the combination of contextualized and

decontextualized testing is suitable to assess the *level of mastery* of the form-meaning link (see also Laufer, Elder, Hill, & Congdon, 2004; Schmitt, 2010). In their experiment, the contextualized test and decontextualized recall (i.e., L2-L1 translation) and recognition (multiple choice) tests of form-meaning association were conducted as measures of three different levels of mastery. These two types of decontextualized tests, recall and recognition tests, have often been applied in research focusing on contextualized vocabulary learning (e.g., Pulido, 2007; Sonbul & Schmitt, 2010; Waring & Takaki, 2003; Webb, 2008), because researchers can assess partial knowledge of word meanings using these tests (Nation & Webb, 2011).

The present study focuses on a learning mode that is (a) intentional, (b) contextualized, and (c) translation-based. Through Experiments 1 to 6, Japanese-speaking students are asked to learn a list of new words such as *clutch*, presented with their translations, in this case *gui-to nigiru*, and example sentences such as, "The festival was extremely crowded, so Emily clutched her father's hand very tightly to keep from getting lost." Hereafter, the term intentional learning refers to a deliberate memorization task to learn such a list within a certain time limit. Recent studies indicated that this mode of learning is effective in obtaining not only explicit knowledge but also implicit knowledge (Elgort, 2011; Elgort & Warren, 2014; Sonbul & Schmitt, 2013). In the present study, the experimenter administered a contextualized recall test after vocabulary learning and a proficiency test either before or after the learning/testing phase. One may argue that this type of learning is uninteresting and more communicative tasks should be enhanced in today's EFL classrooms; this thesis is no objection to that kind of argument. However, the author believes that the research findings from this study focusing on intentional vocabulary learning are also suggestive for other modes of learning (e.g., incidental learning tasks).

2.2.2 Incidental Learning Tasks

Another typical mode of vocabulary learning is the incidental learning. If the source of learning were limited to decontextualized approaches such as rote memorization using translations and a teacher's explanation of word meanings, the number of new words encountered would be quite a bit lower. The effect of contextualized learning was first proposed in studies on L1 vocabulary acquisition, but it has also been regarded as important in the EFL context and studied widely in that regard (e.g., Hunt & Beglar, 2005; Waring & Takaki, 2003). In the 1980s, many researchers indicated that L1 vocabulary knowledge is gained through reading; Nagy et al. (1985) was a major contribution. They hypothesized that incidental learning from free reading is a major source for vocabulary acquisition during the school-aged years. As a result, they tentatively estimated the number of new words the typical middle grade child learns in a year through reading to be between 750 and 5,500, its point-value estimate being 3,125. However, in the case of L2 acquisition, early studies on vocabulary learning only from context exposure through reading indicated a lower gain of knowledge, especially when compared to rote learning using translations. The significant L2 studies are summarized in Table 2.1.

Among these studies, test scores after learning procedures varied, and it is quite difficult to generalize the findings. As Webb (2007a) suggested, it was problematic that most studies of contextualized vocabulary learning did not exclude the variable of learning intentionality. Therefore, to investigate the relation between context reading and vocabulary learning, it is important to distinguish between the two learning modes: incidental and intentional learning. Hereafter, incidental learning refers to the unintentional gain of certain kinds of knowledge or the improvement of memory. In fact, the reported knowledge gain of unfamiliar words was much higher in intentional learning studies such as Prince's (1996)

Table 2.1

Past Studies on Contextualized Vocabulary Learning

Study	Mode	Participants	Context type
Pickering (1982)	INT	26 JH + HS students	Sentential context in L1
Pitts et al. (1989)	INC	74 adult learners	Parts of two narratives
Day et al. (1991)	INC	(1) 191 HS students	A short narrative
	INC	(2) 397 U students	A short narrative
Hulstijn (1992)	INC	(1) 65 adult learners	An expository text
	INC	(2) 98 adult learners	An expository text
	INC	(3) 45 adult learners	An expository text
	INC/INT	(4) 52 JH students	An expository text
	INC	(5) 35 adult learners	An expository text
Dupuy & Krashen (1993)	INC	42 U students	The sequence of a video
Prince (1996)	INT	48 U students	Sentential context in L2
Laufer & Shmueli (1997)	INC/INT	128 HS students	Sentential context in L2/ Context from a textbook
Horst et al. (1998)	INC	34 U students	A whole narrative story
Waring & Takaki (2003)	INC	15 U students	A whole graded reader
Webb (2007a)	INT	84 U students	Sentential context in L2
Webb (2007b)	INC	121 U students	Sentential context in L2
Webb (2008)	INC	50 U students	Sentential context in L2

Note. INC = incidental learning; INT = intentional learning. JH = junior high school; HS = high school; U = university.

study than in incidental learning studies such as Pitts, White, and Krashen's (1989) study, the successful learning rates being 48–51% versus 6–8%, respectively. Sonbul and Schmitt (2010) found that test scores were higher after an explicit, intentional task than after an implicit, incidental task. This is why incidental learning through context reading has been considered "inefficient." Schmitt (2010) defines incidental learning as "learning which accrues as a by-product of language usage, without the intended purpose of learning a particular linguistic knowledge" (p. 29).

However, the rates of vocabulary learning were quite different even among incidental learning studies and among intentional learning studies. As Webb (2008) argues, another reason for the variation in test scores lies in the feature of context. In his study, the textual factor concerning the amount of information in the context (hereafter *informativeness*) was taken into consideration. As a result, lexical items were learned better in more informative contexts than less informative ones. This finding strongly suggests that textual features should be taken into account at least in research on incidental vocabulary learning. Note that researchers studying lexical access in the mind often call this factor the context constraint (e.g., Griffin & Bock, 1998; Schwartz & Kroll, 2006; van Assche, Drieghe, Duyck, Welvaert, & Hartsuiker, 2011; van Hell & de Groot, 2008). They assume that this factor affects the speed and timing of semantic activation in the mind when reading a context. However, this assumption is seemingly contradictory to the results of other researchers' experiments such as Joe (2010) and Zahar, Cobb, and Spada (2001). In their studies, the effect of context richness had no significant effect on vocabulary knowledge gain.

Unfortunately, as Table 2.1 shows, researchers have rarely included a careful comparison of textual features in examining the learning rate through reading. In addition, researchers tended to use different learning materials, making it more difficult to generalize the findings from each study. For example, researchers used narrative stories (Day, Omura, &

Hiramatsu, 1991; Pitts et al., 1989; Waring & Takaki, 2003), expository texts (Hulstijn, 1992), sentential context (Laufer & Shmueli, 1997; Webb, 2008) and a combination of written and spoken contexts (Dupuy & Krashen, 1993). Among prior studies, Webb's (2008) materials seem to be sophisticated due to the following three facts. First, the context feature was assessed in terms of informativeness. Therefore, researchers could easily take this factor into consideration. Second, all the target words were pseudowords in disguised forms. This procedure eliminated the possibility that learners had prior knowledge of the target lexical items. On the other hand, one may argue that the use of pseudowords does not reflect the real learning experience. Therefore, the present study conducted some experiments with pseudowords and others with real words. Third, each context displayed an appropriate level of difficulty; the researcher confirmed through a pilot study that no contexts surrounding the target words had words that were unfamiliar to university students. For these reasons, the current study first used Webb's materials, and the features of the context were further examined.

To reveal how the process of context reading affects vocabulary knowledge, this study focused on intentional, not incidental, vocabulary learning. In particular, Studies 1 and 2 explored the condition under which deeper comprehension of contexts improves participants' recall test performance; Study 3 further examined in which process participants utilized contextual information for vocabulary learning. However, the research findings should also be relevant to incidental learning tasks such as extensive reading, dictogloss, and story retelling. Dictogloss is a kind of dictation task that requires learners to write down something they have listened to. However, dictogloss is different from typical dictation tasks in that students form pairs or groups to exchange what they have understood. In these tasks (i.e., extensive reading, dictogloss, and story retelling), students' attention is fully or partially devoted to the message that the text conveys, but they are expected to learn the vocabulary used in the text during the

task. It would be interesting to investigate whether the effect of the context types the present study focused on can be observed across different activities. This matter is further discussed in Chapter 7.

2.3 Process of Context Comprehension

2.3.1 Levels of Mental Representation

As in L1 reading, text comprehension in an L2 includes different aspects such as *lower* and *higher level processing* (Grabe, 2009). Lower level text processing is characterized in terms of recognition of graphical features and analysis of what is explicitly written in the text. For example, cognitive processes such as recognition of letters and words, access to mental lexicon, and syntactic parsing of a sentence are regarded as the initial stage of reading comprehension (i.e., lower level processing). However, to achieve deeper understanding of the author's message or a situation described in the text, readers often need to activate their prior knowledge and make inferences about contextually relevant information. Through this kind of text processing, or the higher level processing, readers integrate pieces of textual information together with their background knowledge in their mind. L1 readers and advanced L2 learners can comprehend a text very fast and accurately because their lower and higher level processing are quite efficient, or even automatized, to some extent. However, it is not the case for beginner learners. It is often difficult for L2 readers to understand a text fluently by making appropriate inferences because their cognitive resource is devoted to the lower level text processing, which is not fully automatized. Therefore, in L2 reading classrooms, it is advisable for teachers to monitor carefully if their students are achieving deep understanding of the reading materials or struggling with lower level processing.

In spite of the suggested close relation between reading and vocabulary acquisition (e.g., Schmitt, Jiang, & Grabe, 2011), the mental process of reading has rarely been considered in vocabulary research. Among reading researchers, it is widely admitted that there are different

levels of comprehension. The most well known discrimination may be the levels of *surface structure*, *propositional textbase*, and *situation models* (van Dijk & Kintsch, 1983; see also Grabe, 2009; Tapiero, 2007). The first level, surface structure, is the representation of the verbatim input; the second level, propositional textbase, is related to readers' propositional understanding of a text. These two levels of mental representation are based on text information rather than readers' individual experience or background knowledge; Tapiero (2007) called these two levels *semantic representation*. The third level, situation models, includes readers' interpretation of a text and its depiction in their mind. According to Tapiero, the notion of the situation model is useful in explaining why a discourse may be viewed from different points of view among readers. Discrimination of the situation models representation from the other two types of representations is important because the higher the level of representation readers construct, the more information they can retain in their memory.

In van Dijk and Kintsch's (1983) model, the situation model is regarded as the most heavily interrelated and integrated structure. In their models, the situation model is related to the highest structure in a reader's working memory. Using this model, van Dijk and Kintsch presented empirical data showing that mental representations at the situation model level are remembered better than those at other levels of representation. This assumption may be consistent with our intuition that learners can remember texts better when they can imagine the described situation well. For example, readers may arouse mental imagery of the situation when they read the following sentence, adapted from Webb (2008): *Your brother wrote his name on this paper*. In contrast, the next sentence may evoke less imagery: *Perhaps this happened. Or perhaps she's ill and can't remember who she is*. The difference between the two contexts can be explained in terms of the ease of constructing the situation model. However, as far as the present author is aware, no past studies have examined whether the ease of constructing mental representations affects vocabulary knowledge gain in the intentional mode of learning.

This study hypothesized that when learners read example sentences for vocabulary learning, the comprehension process is somewhat different for more versus less imageable (or meaningful) contexts. If this is the case, there might be some cases in which learners' deeper understanding of the meaning of the context promotes understanding of the target word meaning or memory storage of what they have learned (e.g., L1-L2 link). However, it is unnatural to suppose that context reading for vocabulary learning involves the same cognitive process as it does for reading a novel for pleasure or skimming a newspaper. Readers' attention should be paid differently, and reading accuracy and speed should also differ when reading a novel, a newspaper, or a list of example sentences because the reading purposes are different. The next section overviews the general concept of reading purpose and some expected characteristics of context reading for vocabulary learning.

2.3.2 Types of Reading Purpose

Reading accuracy and speed change according to the purpose for reading. Carver (1992) proposed five different levels (or *gears* in his terminology) of reading: scanning (fastest), skimming, rauding, reading to learn, and reading to memorize (slowest). The most basic one is rauding, which refers to reading for general understanding; fluent readers can do this at about 300 words per minute (wpm). For example, reading a novel for fun is sometimes rauding, but it can also be faster than that, especially in L1 reading. Compared to this mode, reading for learning and memorizing a text are generally slower. Carver estimates the rate of reading to learn at about 200 wpm and reading to memorize at about 130 wpm, which allows readers to rehearse what they have just read. However, the reading rates differ across readers.

Applying Carver's framework to intentional learning of L2 vocabulary, reading example sentences can be regarded as reading to learn; otherwise, learners read contexts to memorize. However, the mental process of context reading for vocabulary learning seems different from that of normal reading; learners' attention should be devoted to new words

embedded in the context rather than the meaning of the context itself. Grabe (2009) also identified six academic purposes for reading: (a) reading to search for information, (b) reading for quick understanding, (c) reading to learn, (d) reading to integrate information, (e) reading to evaluate, critique, and use information, and (f) reading for general comprehension. Based on his framework, it becomes even clearer that context reading for vocabulary learning does not require a highly complicated process such as integration or evaluation of information. In addition, questions given to readers after they have finished reading have a great effect on their reading performance (Alderson, 2000). For example, questions concerning higher-level information about the text content result in good retention of both factual and inferential information. This suggests that reading comprehension and text memory are enhanced when readers pay a lot of attention to the text content to elaborate their understanding. In other words, when readers' attention is devoted to memorizing new words, comprehension of contexts might be suppressed, or at least undergone within a limited cognitive resource.

Studies on incidental vocabulary learning through reading also suggested that reading purpose influences the probability and degree of knowledge gain. In Swanborn and de Glopper (2002), three reading purposes were compared in terms of L1 readers' incidental learning of new words while reading a 1,500-word informative text. The three purposes were (a) reading for learning about the topic, (b) reading for comprehension of the whole text, and (c) free reading, which did not direct readers' attention to any specific features of the text. Reading for text comprehension was different from reading for learning the topic in that comprehension of the whole text required both understanding of what the text was about and comprehension of topic-related words. The results showed different effects of reading purpose across proficiency groups. Low-ability readers hardly learned new words regardless of the purpose. In contrast, high-ability readers gained knowledge even under the free reading condition; however, their test scores were better under the text comprehension condition.

Interestingly, reading for learning about the topic did not result in significant gain. Although Swanborn and de Glopper regarded the small number of participants as the first reason for the lack of significance, they also considered the effect of selective attention among the high-ability readers. Namely, more skilled readers were more capable of estimating which words were important; in other words, they knew which words could be skipped. Direction of selective attention seems to be a key factor not only for L1 but also for L2 readers.

2.3.3 Context Reading for Vocabulary Learning

Unfortunately, as far as the author is aware, no theoretical framework directly predicts how learners distribute their cognitive resources while they are memorizing a list of L2 word forms, L1 translations, and L2 contexts. Although there have been a number of studies and discussions on contextualized learning of L2 vocabulary (e.g., Wesche & Paribakht, 2010), most of them have focused on incidental or guessing-based vocabulary learning (see Figure 2.4 in Section 2.2.1). These studies often utilized think-aloud methods (e.g., Rott, 2005) and eye-tracking methods (e.g., Chaffin, Morris, & Seely, 2001; Godfroid, Boers, & Housen, 2013). Applying the discussion in Section 2.3.2 to L2 vocabulary learning, one may argue that context reading in intentional vocabulary learning is useless because learners' attention would be distracted by the context presentation. This section mentions two different theories that explain how learners process and memorize presented items; the first one supports the *context distraction view* above, whereas the second provides some counterarguments.

First, Barcroft's (2002, 2003, 2004) theory, called the *Type of Processing-Resource Allocation* (TOPRA) *model* seems to be applicable to this study. The general form of the TOPRA model can be illustrated as Figure 2.5; however, the present author added some features to the original model. This model was designed to explain how it is difficult for learners to process different types of input, such as L2 word forms and their meanings. This

model assumes a sort of tradeoff between one type of input and another; if a learner processes input more for Type A, the line between Types A and B in the model moves to the right, decreasing processing Type B. The model further assumes that the amount of input processing predicts the success of learning. For example, if a learner paid more attention to the target word forms in intentional vocabulary learning using a word list, other elements such as word meanings represented as translations and contexts would be processed less, resulting in more orthographic or morphological learning but less semantic learning. The thicker outer lines in the model remain stable, representing the restricted amount of processing resources available to a learner. This resource seems to be closely related to the learner's working memory capacity and proficiency in the target language.

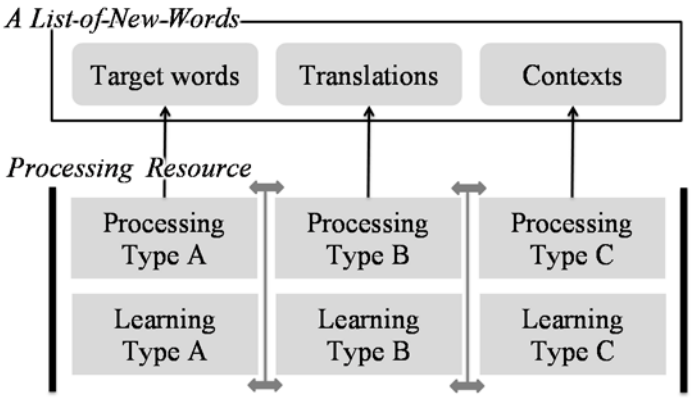

Figure 2.5. The general form of the TOPRA model, adapted with some modifications from Barcroft (2003, p. 549). The present author added a list of to-be-learned words paired with translations and contexts.

According to this model, it is expected that if the learner is not familiar with the word forms in a presented list, presenting example sentences is not suggested because it might distract learners' attention and consequently inhibit learning of word forms. Otherwise, learners might ignore the contexts if their restricted amount of processing resources is not

sufficient for processing contexts. This seems compatible with Webb's (2007a) result showing that the presentation of contexts did not enhance translation-based vocabulary learning. However, one may argue that this view is too simplistic and the effect of context reading should be assessed more carefully. For example, as reviewed in Section 2.1.3, there are both gaps and overlaps between L1 and L2 translation pairs, and learners need various contextual inputs to learn the semantic features of the L2 word incrementally. If learners read contexts to understand what the target word refers to in a specific situation, the process involved in context reading can be seen as an elaboration of meaning. As predicted in *Involvement Load Hypothesis* (Laufer & Hulstijn, 2001; see also Hulstijn & Laufer 2001), which originally focuses on incidental vocabulary learning, a more elaborated process (i.e., high involvement load) leads to more successful learning of new words.

In fact, Hasegawa (2010, 2013) found that L2 learners could depict rich mental representations when the context was provided in addition to pairs of L2 word forms and L1 translations. His results seem to be compatible with Murnane, Phelps, and Malmberg's (1999) theory, called the *Item Context Ensemble* (ICE) *model*. This model assumes that traces in learner memory can consist of three types of information: an item, an associated context, and an ensemble. Here, the term *ensemble* represents the integration of the to-be-remembered item and the context presented in the learning phase. According to this model, when an item is learned in relation to a context, the relationship between the item and context may become meaningful, resulting in better retention and recall of the item. Their theory was applied to L2 vocabulary learning context in Franco-Watkins and Dougherty (2006). In their experiments, encoding instruction was given to participants asking them to form an interactive imagery between the item and the context such as a mental picture or a background. For example, if the target item is *apple* and the context describes a school bus, a participant might raise a mental image of "large red apples seated on the bus" (p. 974). Based on this theory, even in

the intentional mode of learning, context presentation might enhance learning of new words if the learners can form a mental image that integrates the new word and the context for it. To understand how opposing predictions could be proposed concerning the effect of context (i.e., distraction and integration views), it is essential to review past studies on mental imagery during reading and learning.

2.4 Generation of Mental Imagery

2.4.1 Reading Comprehension and Mental Imagery

The mental process of reading has been a point of interest in psycholinguistics. As mentioned in Section 2.3.1, the elaborateness of reading comprehension is explained in Walter Kintsch and his colleague's theoretical work (e.g., van Dijk & Kintsch, 1983) and mentioned in other mental models such as *Construction-Integration Model*, *Structure Building Framework*, and *Landscape Model of Reading* (see Grabe, 2009). What is widely acknowledged among researchers of reading is that there are different levels of comprehension. The most well known differences may be between *surface structure*, *propositional textbase*, and *situation model* (Tapiero, 2007; van Dijk & Kintsch, 1983), in which the first level, surface structure, consists of the representation of the verbatim input; the second level, propositional textbase, relates to readers' propositional understanding of the text; and the third level, situation models, includes readers' interpretations and mental depictions of the text. The reading process is most elaborate at the situation model level.

When an experimenter would like a reader to construct an elaborated representation in his or her mind, it is sometimes quite useful to give the reader instructions on how to read the text. A typical instruction may be to formulate a mental image of the text. In previous studies on reading comprehension, this kind of *imagery instruction* has long been used to help a reader construct highly elaborated representations (e.g., Alba, 1984; Horiba, 2002). In fact,

less skilled readers can understand both explicit and implicit information in a text when undergoing imagery instruction (Gambrell & Jawitz, 1993). Also, the instruction to have participants rate the vividness of a mental picture or image they constructed from a story has been a traditional method in incidental learning experiments. For example, Thorndyke (1977), who is a renowned advocate of the Story Grammar theory, conducted an experiment using such a method (p. 100). A clearer example showing the effectiveness of imagery instruction is a study on reading for problem solving. Antonietti (1991, 1999) showed that mental imagery was quite useful in logical reasoning while reading. Without the help of mental imagery, readers would have difficulty solving a problem such as the one described here: "A rope ladder was hanging from a boat so that the ladder had six rungs above the sea. The distance between any two rungs was 30 cm. At high tide, the sea level rose 70 cm. How many rungs were above the sea at high tide?" (Antonietti, 1999, p. 412). All these studies indicate the effectiveness of intentional generation of mental representations. However, the construction of an elaborate mental image like a situation model might amount to effortful processing (Zwaan & Radvanski, 1998); therefore, it is quite possible that such instruction may interrupt the reading flow if the reader does not have enough cognitive sources available, or if the text is too difficult to arouse an image.

The present study investigates the effect of mental imagery generation on contextualized vocabulary learning, taking into account the contextual ease of evocation of a mental image. The psycholinguistic factor of ease of evoking mental imagery from linguistic materials is generally called *imageability* (e.g., N. Ellis & Beaton, 1993b; de Groot, 2011; Prado & Ullman, 2009; Steinel, Hulstijn, & Steinel, 2007; Woollams, 2005). This concept can be adapted to any kind of linguistic material from a word to a text. Past studies on L1 and L2 reading showed that texts with higher imageability were much easier to retain in the mind (Duthie, Nippold, Billow, & Mansfield, 2008; Gambrell & Jawitz, 1993; Krasny & Sadoski,

2008; Sadoski, Goetz, & Fritz, 1993; Sadoski, Goetz, & Rodriguez, 2000). Imageability is known to highly correlate with *concreteness*, which indicates whether linguistic material refers to physical entities that can be perceived by the senses (e.g., de Groot, 2011; Ogawa & Inamura, 1974). However, de Groot (2011) states that the concrete/abstract words in her studies were generally derived from word imageability norms, not concreteness norms (p. 114). As another example, Davey's (1988) ratings on concreteness were done in terms of "ease of imageability." Therefore, the present study adopts imageability norms rather than concreteness norms. Table 2.2 summarizes the terms concerning the features of the material that have been mentioned in this paper (see also Section 6.1.1). In most cases, imageability of linguistic materials was rated by experimental participants using a 7-point scale. This procedure was also carried out in the present study.

Table 2.2

Brief Explanation of Imageability, Concreteness, Informativeness, and Constraint

Terms	Explanation	References
Imageability	This indicates the ease to evoke of evoking mental imagery from linguistic materials.	Steinel et al. (2007)
Concreteness	This indicates whether linguistic materials refer to physical entities that can be perceived by the senses.	de Groot (2011)
Informativeness	This indicates the amount of context information useful in guessing an unfamiliar word.	Webb (2008)
Constraint	This indicates the degree to which the sentence frame preceding the target word biased that word.	Schwartz & Kroll (2006)

Note. Imageability and concreteness are closely related, as are informativeness and constraint. Differences between the imageability and informativeness are discussed in Section 6.1.1.

Researchers have found that imageability of a sentence is affected by how many words in that sentence are concrete and imageable. For example, in Begg and Paivio (1969), average imageability ratings of nouns included in the 50 more imageable (i.e., concrete) and 50 less imageable (i.e., abstract) sentences were quite different (6.27 and 3.18, respectively; the rating was based on 7-point Likert scale). According to them, a concrete sentence such as "The fat boy kicked a girl" (p. 821) can be represented in a reader's mind as an action picture, where the meaning of the entire sentence forms an organized unit (i.e., mental imagery). Therefore, when learners read an imageable sentence, they tend to retain the overall meaning; on the other hand, when they read a less imageable sentence, they tend to remember the superficial wordings (see also Rowe, Schurr, & Meisinger, 1978). With respect to the materials used in the present study, it might be worth noting that the more imageable contexts tended to include more concrete nouns than the less imageable contexts. When the keyword abstraction analysis was conducted using AntConc3.2.4w (http://www.laurenceanthony.net/antconc_index.html), the top 20 keywords of the more imageable contexts were *the, I, London, my, a, can, chef, dark, doctor, hand, into, last, left, let's, long, lot, new, our, restaurant, station* and those of the less imageable contexts were *they, that, but, she, about, did, his, it, had, saw, the, in, her, big, for, he, large, many, not, some* (see also Section 6.1.1). The present study did not used sentences that requires learners of mental rotation and other unnatural types of context because it seemed quite strange to use them as example sentences of new words (e.g., "The letter *W* rotated 180 degrees and written in lower case, looks like the letter *m*"; Just, Newman, Keller, McEleney, & Carprnter, 2004, p. 115).

In general, as mentioned earlier, more imageable texts have proven to be easier to remember. In Sadoski et al. (2000), text imageability was the strongest factor affecting learners' memory of the material in the text. The impact of the imageability effect was different among text types (i.e., persuasion, exposition, literary, and narrative texts). In their

study, memory of the text was examined with a cued recall task, the cue being the title of each text. In addition, further research might be needed regarding the ways that learners of different developmental levels generate and utilize mental imagery from context. In a recent experiment with L1 readers by Duthie et al. (2008), the reported mental imagery differed according to the reader groups. Because of both language proficiency and richer background knowledge and experience, the elder readers generated more appropriate imagery. Although it may be a shortcoming to expect the same results between L1 and L2 reading, there should be some similarity between these groups (see Krasny & Sadoski, 2008). For example, a context such as *Mrs. Smith was a very pious woman who attended church services regularly* would be an imageable context for L2 learners who have background knowledge about church services. In contrast, if learners have no knowledge about religion or church, the same context would have meager imageability. Thus, the interaction effect between context imageability and learner factors such as proficiency and background knowledge should be further examined in future research. However, comprehension and retention of context has been investigated exclusively in reading research; this notion has been rarely considered in vocabulary research. The current study should be highly suggestive because it examines whether the findings from reading research can be applied to vocabulary learning.

2.4.2 Vocabulary Learning and Mental Imagery

Research on mental imagery has been carried out for imageability more as a lexical than a contextual feature. L2 vocabulary learning can be affected by many lexical properties including word imageability. N. Ellis and Beaton (1993b) introduced 10 major psycholinguistic variables relating to vocabulary learnability. For example, the greater the overlap between the phonological features of L2 and L1 words, the easier it is to learn that L2 word. Other examples that have received a lot of attention in L2 learning research are word

frequency and length. Learners can more easily process and memorize L2 words occurring with a higher frequency or words containing fewer letters than words with a lower frequency or more letters. Among such variables, they described imageability as the degree to which a word arouses a mental image. Word imageability is often categorized as a semantic or meaning property, together with word concreteness, ambiguity, and meaningfulness (de Groot & Poot, 1997; Woollams, 2005; Yokokawa, 2006). Among such related lexical properties, imageability has been revealed to have a more direct influence on conceptual access than concreteness (Richardson, 2002) and ambiguity (Woollams, 2005). According to Woollams (2005), whereas both lexical ambiguity and imageability have facilitating effects on word naming, only imageability is affected by experimental manipulation to detect semantic characteristics. This result has proved the semantic locus of the imageability effect.

There are two methods of obtaining imageability values for experimental or educational purposes: (a) to collect imageability ratings from participants, and (b) to use a psycholinguistic database. First, researchers may ask participants to rate the degree to which each linguistic item produces a mental image, using a 7-point scale (e.g., de Groot & Poot, 1997; Steinel et al., 2007). For example, the imageability value of the Dutch word *stad*, which means "town" in English, is 6.04 in de Groot and Poot (1997); that of the English idiom *to paint the town red* is 4.60 in Steinel et al. (2007). The other way to obtain the imageability value is to use a published dataset (e.g., Woollams, 2005). As introduced by Yokokawa (2006), the most common and reliable database of the lexical properties of English words is the MRC psycholinguistic database (Coltheart, 1981). The MRC database provides 26 kinds of lexical properties of English words, including imageability. For Japanese words, on the other hand, the NTT database offers imageability norms that have been commonly used for psycholinguistic research (Sakuma et al., 2005). However, it should be noted that words with a high imageability value in the database are not always imageable for L2 learners. It is necessary to consider participants' characteristics and other factors affecting imageability in

imagery research.

As discussed earlier, word imageability as a lexical property has some features in common with other lexical characteristics. However, the word imageability shows one clear distinction: its susceptibility to contextual conditioning (Denis, 1989). For instance, in a sentence such as *He cut down a tree with a tool*, the word *tool* is more imageable than in *Tools are necessary for technological progress*. Concerning this point, Paivio (1991) has pointed out that imagery is determined by the following three variables: (a) the image-evoking value of words, (b) interference of experimental procedures, and (c) individual differences in imagery abilities. The second variable (b) covers the contextual factor. In other words, the factors affecting imageability include conceptual concreteness, linguistic context, and individual capacity.

Many studies have examined mental imagery generation as related to lexical items rather than to contexts. In particular, such studies focused on a strategy to help learners with L2 vocabulary, called the *keyword method* (e.g., Avila & Sadoski, 1996; N. Ellis & Beaton, 1993a, 1993b). For example, by using this strategy, English-speaking learners can associate the German word *Sperre* with its English translation *barrier* by using the mediating image of "SPARROW in a station BARRIER" (N. Ellis & Beaton, 1993a, p. 557). N. Ellis and Beaton (1993b) examined many psycholinguistic determinants of L2 vocabulary learning and regarded imageability as a strong determinant of learnability. Furthermore, Hasegawa (2010) reported that after a mixed task of contextualized and decontextualized learning of unfamiliar L2 words, the posttest scores were higher for more imageable items than for less imageable ones. In addition, the imageability of the target items increased when a context was available.

However, it is also worth noting that learners with lower proficiency might be poor at utilizing imagery representation for vocabulary learning. According to a survey by Mizumoto and Takeuchi (2009) regarding the instruction of vocabulary learning strategies, imagery strategies were found to be difficult for them to learn through explicit instruction. Five of the

strategies investigated were as follows: (a) making a mental picture of items associated with a word's meaning, (b) linking the learner's own personal experiences to the word, (c) creating an image of the spellings or orthographic forms, (d) using the keyword method (keyword mnemonic technique), and (e) imagining whether the word meaning is negative or positive. The results of their experiment showed that explicit instructions were effective for developing some groups of learning strategies but not mental images. Students' use of the following strategies improved after instruction on input-seeking strategies (e.g., exposing themselves to English vocabulary by extensive reading), oral rehearsal strategies (e.g., saying a word aloud repeatedly), and association strategies (e.g., associating a word with its synonyms or antonyms they already know). Mizumoto and Takeuchi analyzed that the use of imagery strategies was not increased after a series of instruction because the strategies were difficult for them to use, even if the strategies appeared useful. One participant reported "I cannot think of a mental image or mnemonics for the target vocabulary. I'm bad at making them by myself. Also, imagery or mnemonics are not suitable for all the words. I'd rather spend my time on writing or vocalizing the target words" (p. 443).

Nevertheless, as explained earlier, it has been empirically and theoretically shown that vocabulary learning is dramatically enhanced by the imageability of the concept. Furthermore, as Hasegawa (2010) indicated, imageability of the lexical items can be increased by context presentation. It was indicated that imagery representation can be constructed through reading a context and that the context can make the lexical items more imageable. It is therefore quite plausible that the mental representation aroused from the context has something to do with learning the lexical items in the context. However, such a possibility was not in the vision of most preceding studies on contextualized vocabulary learning. Thus, the present study focuses on the mental representation of the learning contexts in relation to the target words.

2.4.3 Context Imageability Effect

In spite of the potential difficulties mentioned in Sections 1.1 and 2.2.1, contextualized learning is often regarded as the best way of learning vocabulary (see Prince, 1996). The view that EFL vocabulary should be learned in context emerges from two general observations. The first is that contextualized vocabulary learning enables learners to encounter both frequent and infrequent words repeatedly (Nation, 2009). The second reason for the importance of contextualized learning is rooted in the theory of mental lexicon development, which is more directly related to the present study's focus than word frequency is. According to Jiang (2000), EFL words learned through rote memorization using translations cannot reach the "final level" of lexical development (see Figure 2.2 in Section 2.1.3), where the semantic, syntactic, morphological, and formal representations of vocabulary are integrated into the lexical entry.

The potential differences between contextualized and decontextualized learning are exemplified in Figure 2.6. Although the figure is a little simplistic as a representation of this complicated theory, it is nevertheless useful as a starting point to help us discuss the influence of the learning method on mental lexicon development. The model suggests that if, for example, a Japanese student learns English vocabulary without context, his or her lexical development will not completely close the gap between the L1 and L2, and as a result, his or her lexical knowledge will be a compound made up of English and Japanese lexical representations. In contrast, contextualized learning is assumed to fill in a lexical entry for a particular learner by associating a conceptual representation with a word-form representation. However, this facilitating effect assumes that the learner can generate an elaborate representation in his or her mind from the reading context.

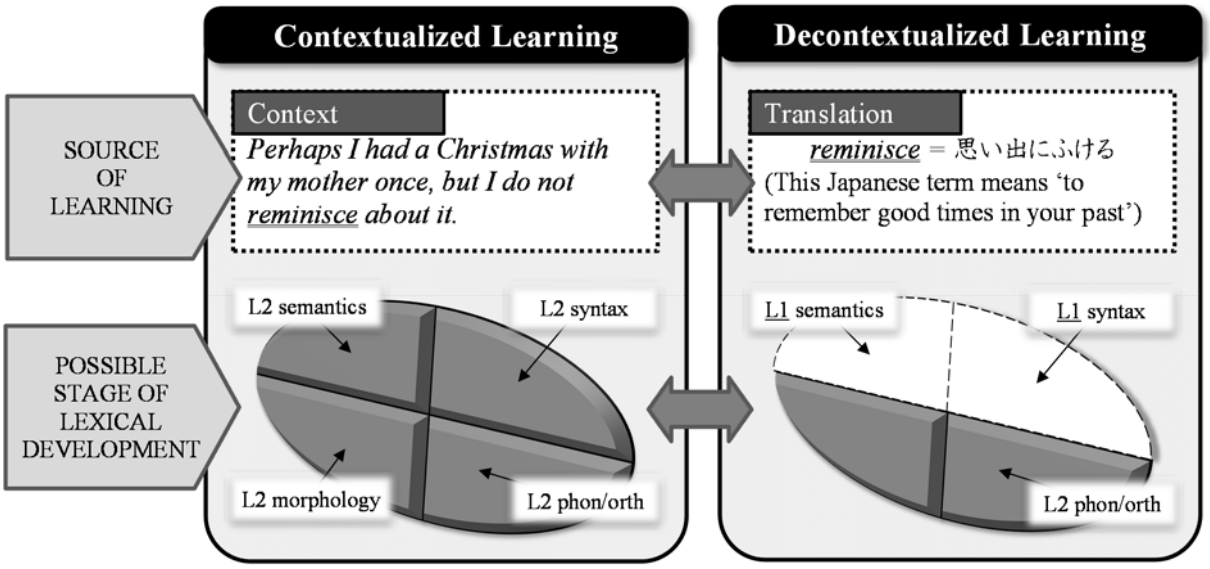

Figure 2.6. Plausible fruits of the two approaches, adapted from Jiang (2000, pp. 53–55), with slight modifications. The example sentence was developed based on Webb (2008, p. 243), replacing a pseudoword with a real, low-frequency word. The terms *phon* and *orth* represent *phonology* and *orthography*, respectively.

Based on the research findings that more imageable contexts that specify a concrete situation or remind learners of related memories can be understood and learned better than less imageable ones (see Section 2.4.1), one might expect that the effect of context presentation in intentional vocabulary learning would be different between two word lists, one with easily imageable contexts and one with contexts that are less imageable. Unfortunately, past studies did not examine what would happen if a learning condition were fully arranged to enhance the learner's elaborated comprehension of contexts. If one's mental representation from contexts affects one's representation of word concepts in some way, or provides learners with opportunities to integrate the item and the context (see Section 2.3.3), some effects should be observed in the learners' memory of the new words.

According to *Dual Coding Theory* (Paivio, 1986, 1991) and *Bilingual Dual Coding Theory* (Paivio, 2006), imageable contexts are processed by two mechanisms in a reader's

mind: the verbal and nonverbal (imagery) systems (see Section 2.4.2). In contrast, less imageable ones are likely to be processed dominantly by the verbal system. The better understanding, memorizing, and retrieving of imageable words and contexts are explained in terms of the effective activation of dual processing. For example, when a learner encounters a context such as *Mrs. Smith was a very pious woman who attended church services regularly*, the processed verbal stimuli activate the units in the verbal system called *logogens*. When s/he creates mental imagery from this context, the corresponding units in the nonverbal imagery system called *imagens* are activated through the connections between the verbal and nonverbal systems.

Based on the previous theories and empirical findings, Kadota (2006, pp. 192–194) proposed a working model focusing on how Japanese EFL learners comprehend texts written in English. His model assumed two separated but closely related components: the analytic mechanism, which undergoes phonological and syntactic parsing and other basic cognitive processes, and the global mechanism, which regards more imagery-based or top-down processes of understanding. According to this model, the global mechanism contributes to learners' more rapid and fluent recognition and comprehension in reading. One might suspect whether processing based on visual imagery generated from a context can be separated from a more verbatim comprehension process. However, this assumption was supported by past studies: For example, Fincher-Kiefer and D'Agostino (2004) revealed that interfering with readers' visual imagery inhibited them from gaining a deep understanding of a text (i.e., an inferential process). Based on these frameworks and other theories reviewed in Section 2.3.1, at least two aspects can be assumed with respect to the context reading processes: (a) analysis of a proposition by decoding orthographical, phonological, lexical, and syntactic features and (b) predictive and elaborative processes activating visual, acoustic, and other types of images. However, it was unclear whether these two types of processing, especially the second one, are

carried out while reading example sentences for vocabulary learning; as predicted in the TOPRA model, it might be more economical for some learners to ignore the example sentences and to concentrate on word forms and meanings.

In contrast to such a prediction, according to the author's previous experiments, context presentation affects at least university-level learners' impressions about the target words (see Section 2.3.3). As reviewed earlier, more imageable words are more easily learned in general (see Section 2.4.2). Taking all these factors into account, the effect of context presentation on intentional vocabulary learning should be observed if the following two conditions are satisfied. First, learners should have good lexical knowledge and reading skills that enable them to understand the given context with a certain level of fluency. According to Hasegawa (2013), university students seem more suitable for examining the context effect than younger learners. In his study, high school students in his study did not show any clear context effects on the impression of a list of new words in terms of how easily they can imagine or visualize these word concepts or referents (i.e., word imageability). Second, the contexts should be highly imageable for the learners, enabling them to construct a rich mental representation. In addition, teacher instruction that requires learners to read the contexts carefully or try to imagine the situation described in the contexts may enhance the effect of context presentation. If the present study's expectation can be realized, vocabulary test scores after intentional learning would be different in two cases: one that satisfies the two (or three) conditions above and one that does not.

2.5 Significance of the Present Study

2.5.1 Problems in Relation to the Present Study

This section summarizes the literature review through Sections 2.1 to 2.4 and then points out two unsolved matters, which have already been introduced in Section 1.2. First, as

explained in Section 2.1, vocabulary development in L2 is quite different from L1 because typical L2 learners already have a lexical system in mind. Researchers agree that L2 words are often recognized, learned, and recalled with the mediation of translations in their L1. However, because conceptual features are not fully shared between translation pairs in L2 and L1 (see Figure 2.1), and because translation-based approach cannot cover all aspects of both high- and low-frequency words, contextual input also plays an important role in lexical development in L2. In other words, vocabulary learning is an incremental process, which might be the most effective when the context-based learning is combined with the translation-based approach appropriately. As reviewed in Section 2.2, there are many types of vocabulary learning (see Figure 2.4). This study focuses on translation-based intentional learning of L2 vocabulary using example sentences because this seems to be the most economical way of learning new words in a limited time. In this approach, learners read a list of (a) target words in L2 (i.e., word forms), (b) translations in L1 (i.e., word meanings), and (c) example sentences (i.e., short verbal contexts) to memorize them deliberately. However, previous studies indicated that in this mode of learning, context presentation is not effective at all because the primary source of lexical knowledge is the translation pairs (i.e., target words and their translations).

In addition, as overviewed in Section 2.3, context reading is a complicated cognitive process that can be influenced dynamically by readers' purposes. In particular, when leaners are concentrating on memorizing unfamiliar word forms and their meanings, contextual information should be processed within a quite limited amount of cognitive resource (see Figure 2.5). Therefore, it seems plausible that learners are not sensitive to what contexts are presented in the list of the to-be-learned words. On contrary to this assumption, the present author's previous experiment indicated that L2 learners regard concepts of new words easier to represent in mind when the new words are presented in context (Hasegawa, 2010). As introduced in Section 2.4, psycholinguists have demonstrated that the ease of learning words,

phrases, and sentences are influenced by how the items to be learned are imageable for the learners. Therefore, the present study considers the context imageability effects on intentional vocabulary learning. Based on suggestions from previous research, it was hypothesized that the efficacy of vocabulary learning using translation is affected somehow by the context imageability. If given contexts are more imageable, learners might be able to activate the imagery processor proposed in Dual Coding Theory more fluently and effortlessly (Paivio, 2006), resulting in more effective process of context-based learning, or more precisely, construction of lexical memory that is related with contextual information.

Based on the above suggestions and assumptions, there seemed to be two topics that should be further investigated: (a) how the effect of context reading for vocabulary learning can be maximized and (b) in what process learners comprehend context while their intention is oriented toward memorizing lexical items. This study hypothesized that, after learners encounter new words in context, the contextual information is retained in relation to the new words. Some learners retain context by drawing a picture in their minds of the situation described in the context. If this is the case, L2 learners should be more sensitive to the contextual information than what was expected in past studies. To test if the context imageability affects vocabulary learning, this study compared the effect of more and less imageability contexts on vocabulary test scores in Study 1. Furthermore, the *cognitive strategies* during learning might be different between more and less imageable contexts. This study uses the term cognitive to cover a wide range; namely, compared to competence or ability, cognitive strategies are involved in learner's actual performance such as comprehension and production (V. Cook, 1997). In particular, the present study collected verbal report called *think-aloud protocols* about what learners are reading and thinking during intentional vocabulary learning to examine the difference in cognitive strategies for more and less imageable contexts. As explained in Block (1986), researchers often assume that think-aloud protocols provide a direct view of a reader's mental activity. Moreover,

think-aloud data are most informative about the mental process when learners have difficulty in understanding what they are reading; thus, the think-aloud method is the most helpful approach to study cognitive strategies. In this study, the think-aloud method was applied in Study 3.

However, the effect of context imageability was assumed to be different across learner groups; this matter was focused on in Study 2. In fact, previous studies have found that the effect of a contextualized task is quite different between higher and lower proficiency learners. For example, Griffin (1992) showed that contextualized learning was ineffective for lower proficiency students. Among participants aged 11 to 13, two proficiency groups were determined by the school they attended. A contextualized recall test, in which participants were given a learning context as a cue to recall learned word forms, showed that only students from the higher proficiency school performed better on the test when they had learned new words in context than out of context; those from the lower proficiency school showed the opposite tendency. He explained the cause of this difference between the two proficiency groups in terms of elaborative processing. Because contextualized learning requires more elaborated processing of information, the positive effect emerged only when the participants treated the task as a semantic rather than an episodic task. Thus, the lower proficiency learners in his study, who treated the task as episodic, were less able to learn what was written in context. It seems that the lower proficiency learners' focus was predominantly on the new words themselves, and thus that they could not semantically process the context information. Similar results were reported in Prince (1996), who assessed university students' English proficiency according to their TOEFL scores. The average score of the higher proficiency group was 480 and that of the lower proficiency group, 397. Prince conducted several tests, including a contextualized recall task, using new contexts that were unfamiliar to the participants. In the contextualized recall task, neither the lower nor the higher proficiency group showed a significant difference between contextualized and decontextualized learning

conditions. In fact, the higher proficiency group's scores after contextualized learning were better in the contextualized recall than the translation recall. This shows that reading a context can be effective for higher proficiency learners but a waste of effort for those with lower proficiency. Prince presented two possible explanations for this difference. First, the essential cause of the lower proficiency learners' inability to process context fluently might have been in their lack of lexical knowledge. Second, the lower proficiency group might have had trouble with syntactic analysis, which may have hampered their semantic processing. In Baleghizadeh and Shahry (2011), who demonstrated the significant effect of presenting multiple contexts on vocabulary learning, the participants' proficiency was closer to the advanced group in Prince (1996) according to their TOEFL scores. Again, it was difficult for learners with lower proficiency to process contextual content elaborately or fluently when their attention was mainly on the unfamiliar words themselves.

Research on contextualized vocabulary learning in Japan focusing on the effect of a glossed context has not always take the following two factors into account: the proficiency factor and the testing method. For example, Webb's (2007a) experiment, which was conducted in Japan, did not consider learner proficiency as an independent variable. One exception to this observation is Anezaki's (2003) study, which showed that a lower proficiency group learned the best in a decontextualized, translation-based condition and the least in a contextualized condition. It was interesting that even though Anezaki's participants were beginner learners (junior high school students), the result was similar to those of Griffin (1992) and Prince (1996). However, the higher proficiency group's performance in Anezaki's study was maximized not in the in-sentence condition but in the in-collocation condition. As for testing method, Read and Chapelle (2001) suggested that decontextualized tests have a different purpose from contextualized tests. For example, Nation's (1990) Vocabulary Levels Test is widely used for measuring vocabulary size independent of contexts (called a *context-independent test*). According to Read and Chapelle, the Vocabulary Levels Test

follows the trait definition of vocabulary, which attributes test scores to the abilities and characteristics of the test takers. In other words, tests with the trait definition of vocabulary assume that contextual variables play no significant role in assessing a particular "trait." Another example of a context-independent test is the Vocabulary Knowledge Scale (Wesche & Paribakht, 1996). Even though the highest level learners are expected to provide an English sentence using the target items, this test is a discrete item test and the lexical items themselves are separated from linguistic context. In contrast to these tests, target vocabulary is provided in a specific type of context in *context-dependent tests* such as the Lexical Frequency Profile (Laufer & Nation, 1995), the vocabulary items in section 3 of the Test of English as a Foreign Language or TOEFL (Educational Testing Service, 2006), and the Lexical Density Index for the speaking subtest of *access* (O'Loughlin, 1995, 1997). Such context-dependent vocabulary tests are assumed to have a positive washback effect, encouraging students' learning of words in context (Qian, 2008; Read, 2000). One may suppose that the reason past studies, including Webb (2007a), used context-independent tests was to assess learners' levels of vocabulary knowledge; they aimed to reduce the interference of the word-context associations in the learners' memory. However, construction of such mental associations can be an important process in vocabulary learning, as reported in Schouten-van Parreren (1989). Thus, Webb's experiment should be replicated using context-dependent tests. In summary, the generalization can be made from past studies that less proficient learners might not be very good at utilizing context because they cannot cognitively afford to process contextual content fluently. However, it remains unclear when contextualized learning is the most effective among different group of learners in terms of scores of a contextualized test administered after vocabulary learning. Studies 1, 2, and 3 was designed to investigate the sensitivity to contextual information, the interaction between imageability and proficiency factors, and the cognitive processes of context reading in a step-by-step manner.

2.5.2 Impact in Terms of Pedagogy

The previous section (Section 2.5.1) introduced the two focuses of the current research, namely, the condition that optimizes vocabulary learning using translations and contexts and the mental process involved in the learning process. From these viewpoints, Studies 1, 2, and 3 examined learners' sensitivity to contextual information, relationship between the sensitivity to contextual information and learners' proficiency, and different processes of context reading, respectively. This section briefly explains how this study can provide pedagogical implication concerning vocabulary learning in L2. As Prince (1996) mentioned, there is a widespread view of vocabulary learning which demands the speedy shift from learning words with their translations to relying on L2 context; such a view is shared by many researchers (e.g., Qian, 2008; Read, 2000). However, as repeatedly mentioned in this chapter, it has been unclear how context reading contributes to the learners' lexical development especially word meanings of the target words are already presented explicitly as translations. Although some studies indicated that the contextualized learning of L2 vocabulary is the most enhanced when the translations are given to the targeted unfamiliar words (e.g., Laufer & Shmueli, 1997), others suggested that the contextualization would have no positive effect on learning new words when the translations are already given (e.g., Webb, 2007a). Concerning this matter, Date (2006) argued that vocabulary learning was the most effective when both contexts and translations are presented with the unfamiliar words. The present study aims to reveal how the combined use of translations and contexts can be effective for vocabulary learning.

In Japanese educational context, learners often utilize translation in Japanese when they learn new words and read unfamiliar texts in English, at least until the Course of Study stipulated that high school teachers should mainly use English in the English classes (see Ministry of Education, Culture, Sports, Science and Technology, 2008, 2009). For example, teachers often assign their students homework that requires them to consult an English-Japanese dictionary to learn the meanings of new words used in the textbook. During

using a dictionary, some students might read example sentences and others not. If learner mind was insensitive to contextual information, and if example sentences had little effect on the students' understanding and memory of the new word, example sentence would not be necessary for the learners. Furthermore, it would be even unnecessary to encounter these words again in the textbook. Indeed, these kinds of assumption are not consistent with teacher's belief or researcher's argument; however, unfortunately, there have been no clear evidence showing the effect of context reading on the translation-based learning of new words. Focusing on learners' imagery generation process, the present study explores the function of context reading during the intentional vocabulary learning,

In Japan, English classes in junior and senior high school are under control of the Ministry of Education's new Course of Study guidelines. It might be worth noting that among studies on contextualized vocabulary learning, Japanese students took part in Day et al. (1991), Waring and Takaki (2003), and Webb (2007a, 2008). The new Course of Study stipulates that students should master the lexical items introduced in lower and upper secondary schools through repeated instructions (Ministry of Education, Culture, Sports, Science and Technology, 2008, 2009). In addition, it requires that in order to cultivate communication abilities, teachers should use teaching materials that give sufficient consideration to actual language-use situations and functions of language. In short, the new Course of Study encourages repeated vocabulary learning in meaningful context. Therefore, research is in great demand now on how teachers ensure the fruitful vocabulary learning with the contextualized approach. Thus, a focal point might be the learners' ability in applying the knowledge of instructed words to new context flexibly. Recent studies found that less skilled learners' understanding of low-frequency meanings of homonyms was not as flexible as advanced learners (Ushiro et al., 2010, 2013). Given these results, a typical characteristic of less proficient learner would be that their performance was sometimes successful and sometimes unsuccessful depending on the consistency between

their knowledge and the given context. Concerning this, the present study expected pedagogical implication about the difference in context use between different learner groups.

In spite of these expected impact in terms of pedagogy, one might argue that the vocabulary learning approach that this study focuses on (i.e., deliberate learning of a word list) is not very effective itself because guessing strategies are more important with respect to lexical skills required in L2 reading, after all. Prior research has examined the relationship between reading context and lexical development. According to an early study by Dollerup, Glahn, and Hansen (1989), vocabulary in the reading process is an aggregate of (a) word knowledge, (b) strategies for decoding words, and (c) clues contained by linguistic context. The first component, word knowledge, is particularly required for words used most frequently; it was argued that vocabulary size can be estimated through testing using trait definitions, such as Nation's (1990) Vocabulary Levels Test. It is consistent with our insight that we cannot read a text unless we know the meanings of the most frequently encountered words (see Nation, 2009, 2013). The second component, strategies for decoding words, is required to gain overall comprehension of a text. By using the first two strategies, readers can make use of the third component, clues in linguistic context. Knowledge of less frequently encountered words is not always necessary when reading a text because readers can often decipher meanings of such words from contextual clues. In addition, a new word might be one which an individual reader will encounter only once in a lifetime. Concerning the contextual exposure, Bolger, Balass, Landen, and Perfetti (2008) indicated that the variety of context is a key of knowledge gain. As learners develop their word knowledge, decoding skills, and strategies for using context, they can process a larger number of words effectively while reading a text. Based on this framework, word knowledge gained through intentional learning is essential for the subsequent learning stages (see also Section 2.1.2). Although learners' lexical knowledge cannot be fully developed only through the intentional vocabulary learning, investigation on this learning mode seems to be quite meaningful because learners' attention

to the to-be-learned words is an decisive factor of the efficacy of vocabulary knowledge gain (e.g., Laufer & Hulstijn, 2001).

Chapter 3

Study 1: Effects of Mental Imagery From Context on Vocabulary Learning

3.1 Focus of This Study

Study 1 consists of three main experiments and a pilot study (see Figure 1.1 in Section 1.2). The primary purpose of Study 1 was to examine the effects of mental imagery from context on intentional vocabulary learning. In Experiments 1, 2, and 3, university undergraduates in Japan learned a list of unfamiliar words in English with translations and example sentences and took some tests including a contextualized recall task of the learned words (see Figure 1.2 in Section 1.3). Each experiment compared the effect of more and less imageable contexts on the recall test scores. Based on past studies suggesting little effect of contexts in the intentional vocabulary learning using translations, one might predict that the context imageability would not affect participants' test scores at all. However, if university students can read the imageable context fluently, they might build up mental imagery during reading the context; if this happens, presumably the mental imagery from the contexts would be encoded into the learners' memory and associated with the information about the target word. If the learners were able to take advantage of mental imagery from contexts, the test scores would be higher for the more imageable context condition than the less imageable condition. As far as the present author is aware, no previous research has examined such a context imageability effect on vocabulary learning.

Among the main and pilot experiments, Pilot Study 1 collected data to identify which contexts were more and less imageable than the others were. Experiment 1 examined the context imageability effect on vocabulary learning. Based on the finding from Experiment 1 that the L2 learners are sensitive to context type to some extent, Experiments 2 and 3 replicated the first experiment with some modifications in the experimental design. In particular, Experiment 2 tested the effect of instruction about how to read the contexts;

Experiment 3 examined if it is reasonable to generalize the finding from the former two experiments by comparing the more and less imageable context conditions with no context condition.

3.2 Pilot Study 1

3.2.1 Purpose

The primary purpose of Pilot Study 1 was to collect materials that were appropriate for Experiment 1 and to identify which contexts were more imageable for learners than others. To collect the context imageability data, university undergraduates were presented with a list of contexts written in plain English and asked to rate the ease of evoking mental imagery during reading each context. The secondary purpose was to examine learners' construction and retention of mental association between word and context memories. As mentioned in Section 2.5, it was expected that more imageable contexts were likely to retained in the learners' mind in relation to the new word (i.e., word-context association). Because the process of context reading was discussed in the later section (Section 6.1.1 and 6.3), results of analyses for this secondary purpose were described in detail in Sections 3.2.3 and 3.2.4. In the experiment, university students read 10 glossed contexts including target words that were unfamiliar to them. After reading, they were asked to recall the content of the contexts. In this recall task, the target words were presented as a cue. If the participants incidentally constructed the word-context association in mind, they would recall the context information in relation to the target words. An additional recall task was performed a month later; this delayed task examined how learners lost the word-context associations. Thus, the major interests of this study were in (a) whether and how much learners can construct and retain the memory of context associated with unfamiliar words and (b) whether such word-context associations were affected by context properties, especially the context imageability.

3.2.2 Method

3.2.2.1 Participants

A total of 46 Japanese undergraduate students participated in Pilot Study 1. All of them were first-year students, who were majoring in engineering, sociology, or international studies. They were asked to take part in both the first (i.e., immediate) and the second (i.e., delayed) sessions, but three of them were absent from the second session. The numbers of participants were 46 in the immediate condition and 43 in the delayed.

3.2.2.2 Materials

The target words and contexts, which were selected from Webb (2008), are shown in Table 3.1.

Table 3.1

Target Words and Their Meanings Used in Pilot Study 1 and Experiment 1

Nouns	Verbs
ancon (hospital; 病院), *cader* (lunch; 昼食), *dangy* (street; 街の通り), *hodet* (face; 顔), *masco* (train; 列車), *tasper* (evening; 夕方)	*denent* (remember; ～を覚えている), *faddam* (write; ～を書く), *pacon* (wear; ～を着ている), *sagod* (visit; ～を訪れる)

Note. The meanings are presented in parentheses with Japanese translations; these translations were used as glossary in the reading phase.

As Table 3.1 shows, there were 10 pseudowords, which were disguised forms of existing English words (see also Appendix 3.1). The 10 words consisted of six nouns and four verbs; the ratio of nouns and verbs was determined to approximate their proportional frequency of occurrence in language use (e.g., Kucera & Francis, 1967). The disguised forms were used in order to ensure that the participants had no prior knowledge of the target words.

Each word had two syllables and consisted of five or six letters. In addition, in making the glossary, translation words in Japanese were selected with reference to *Taishukan's Genius English-Japanese Dictionary* (Konishi & Minamide, 2001).

In the experiment, these pseudowords were presented in short glossed contexts consisting of one or two sentences. The numbers of words varied from 6 to 32 (M = 15.60, SD = 6.58). The contexts were taken from the Oxford Bookworm series. Webb (2008) confirmed that all the contexts were not supposed to contain words that were too difficult for Japanese university students. In fact, except for the target words and proper nouns, 92.2% of the running words fell into the first 1,000 frequency band in the *JACET 8000* words list (Committee of revising the JACET basic words, 2003). The first three (i.e., 1,000-3,000) frequency bands covered 98.7% of the running words. The number of contexts was five per target word; there were 50 contexts in total. The factor taken into account in Webb's study was the amount of information related to a target word's meaning. In his study, the contexts were rated on the informativeness by two native speakers and only contexts that were given the same rating by both native speakers were included in the material set. The rating was made on the 4-point scale: *1* and *2* represented *an insufficient amount of information*, *3* represented *a reasonable amount of information but with possibility of misunderstanding*, and *4* represented *a sufficient amount of information for learning* (Webb, 2008, pp. 235–236). As a result, the ratings among the 50 contexts varied between *1* (n = 3), *2* (n = 13), *3* (n = 20), and *4* (n = 14). However, the treatment of contexts that were rated as *3* was problematic in Webb's study; such contexts were treated as either more or less informative according to the situation. To avoid such erratic behavior of contexts, the present study regarded the 16 items rated at *1* or *2* as uninformative contexts and 14 contexts with the rating of *4* as informative. The contexts rated at *3* were excluded from analysis when context informativeness was included as an independent variable. Thus, there were 50 contexts in the analysis of whole recall rates but 30 when the context features (i.e., informativeness and imageability) were

taken into account. The 50 contexts, each of which included one target pseudoword, were randomly grouped into five sets. After the randomization, an analysis of variance (ANOVA) indicated the average informativeness was equivalent between the five sets, $F(4, 25) = 0.32$, $p = .865$, $\eta_p^2 = .048$. Participants got one set each and were asked to read them. In this manner, 10 different target words were given each participant, but the contexts they read were different between participants.

3.2.2.3 Procedure

The participants were asked to take part in two experimental sessions. The procedures are summarized in Table 3.2. The first session included a context reading task and a recall task; the second session included only a recall task. Participants received general instruction before the main experimental tasks; they took a questionnaire at the end of the sessions. The overall procedure was completed within approximately 50 minutes. After the experiment, each participant was asked some general questions about their background. The first session included three tasks: context reading, imageability rating, and immediate recall. Note that the imageability rating was done while reading the contexts. The participants were given a booklet and asked to follow the experimenter's instruction (see Appendix 3.2). The instruction was always provided both orally and visually in Japanese. The participants were first presented with 10 contexts, each of which included one target pseudoword. Because the present study aimed to figure out the memory of contextual information, it was necessary to avoid misinterpretation of target words. Therefore, as in Webb (2007a), each target word was always glossed using L1 translation. During reading, the participants were required to rate the contexts for the ease or difficulty with which they aroused mental images. As in Woollams (2005) and other studies, each imageability rating was made on a 7-point Likert scale. The instruction for this task is presented in detail in Table 3.2.

Table 3.2

Summary of the Procedure in Pilot Study 1

Task	Instruction
	Session 1
Context reading	"Read the English contexts carefully."
Imageability rating	"During reading, rate each context as to the ease to arouse an image. If you can rapidly build up a vivid image, circle *7* on the scale. If you cannot evoke any image from a context, choose *1*."
Immediate recall	"Look at the 10 words below and recall the context in which each word used. Write down anything you remember in Japanese."
	Session 2
Delayed recall	"Look at the 10 words below; they are words you encountered the other day. Recall the context in which each word was used and write down anything you remember in Japanese."

Note. The task instruction was originally given in Japanese.

The imageability data collected here were used in a later analysis. In addition to data collection, the imageability rating task created a situation in which learners concentrated on context reading. No implication was made concerning the recall tasks, and it was assumed that the learning was in the incidental mode. The imageability rating plus reading took approximately eight minutes and all the participants completed it within 10 minutes. After that, a filler task that had nothing to do with the present experiment was performed in 10 minutes. In this filler task, the participants took an English vocabulary test; the data were used for calculating the item reliability and the items regarded as sufficiently reliable were used in Experiment 1.

After the first session, the participants were presented with the 10 pseudowords and

asked to recall the contexts they had read (see Appendix 3.3). They wrote down whatever they could remember in Japanese without reference to the material passages. Whether readers recalled context fully or partially, it can be assumed that the context information was stored in their memories and retrieved using the word-context association in mind. A month later, the second session was held and the same recall task was repeated. This delayed recall task enabled comparison between short- and long-term retention of the context.

3.2.2.4 Scoring

As to the criteria for scoring recall production, the present study adopted Ikeno's (1996) idea unit (IU) analysis. His criteria were based on Carrell's (1991) method and often used in the scoring of recall protocols in Japanese (e.g., Kimura, 1999; Ushiro et al., 2011). All the material contexts were analyzed into IUs by two raters, one was the author and the other was a graduate student who was majoring English education and who was familiar with IU analysis. These raters first independently divided into IUs 20 contexts out of the 50. The interrater reliability was high (percentage agreement = 96.87%). All disagreements were solved through discussion between the raters. After that, the remaining 30 contexts were divided and the interrater reliability was still higher (percentage agreement = 98.23%).

The participants' recall protocols were scored in terms of the production rates of IUs. In Ikeno (1996), each subject received one point for producing an IU when their recall included a verbatim or recognizable paraphrase of the content words. The number of IUs was then transformed into a percentage of the total number of IUs in the original passages, because the number of IUs differed between passages. The current study followed Ikeno's procedure and the production rate of IUs was calculated item by item. To establish valid scoring criteria, the author and a graduate student who was majoring in English education independently scored 20% of the recall protocol data. The interrater reliability was high (percentage agreement = 94.48%). Through discussion between the raters, disagreements were solved by establishing

several criteria such as “Permit a paraphrase of a proper noun into a pronoun” and “Do not score when a participant just copied a translation of the target word.” Based on the criteria, the remaining 70% of the data were analyzed by the author.

3.2.2.5 Analysis

The present study utilizes some statistical analyses such as ANOVA. For all statistical analyses, the alpha level was set at .05. In examining informativeness and imageability effects, these two factors were treated as follows. Firstly, as described in Section 3.2.2.2, context informativeness was adapted from Webb (2008), excluding the moderately informative items (i.e., contexts rated at *3*) from ANOVA. The remaining 30 contexts were further divided into more and less imageable items. Given the mean rating of 5.30 and the median of 5.44, the contexts with more than 5.44 points were regarded as the more imageable contexts. This division was identical with the result of a cluster analysis using the Ward method with the squared Euclidean distance technique. As a result, 18 contexts formed a less imageable group and the other 12 were a more imageable group. Interestingly, the ratings of informativeness and imageability had no significant correlation ($\rho = -.03$, $p = .867$). It enabled the grouping of contexts into four categories: (a) more informative and more imageable ($n = 6$), (b) more informative and less imageable ($n = 8$), (c) less informative and more imageable ($n = 6$), and (d) less informative and less imageable groups ($n = 10$). As in the correlation analysis, a χ^2 test confirmed that the distribution is not biased ($\chi^2 = 0.09$, $p = .765$).

3.2.3 Results

The first criterion was the amount of context information the participants recalled. The data of recall production rates were analyzed in terms of (a) the numbers of recalled contexts per participant and further examined by (b) the percentages of recalled IUs. The average numbers of recalled contexts are shown in Table 3.3, indicating the percentage of the target

words recalled in relation to the context information. As a result, the number of recalled contexts was 4.24 (42.4%) per participant in the immediate condition and 2.04 (20.4%) in the delayed condition.

Table 3.3

The Numbers of Recalled Contexts per Participant in Pilot Study 1

Task	*N*	*M*	*SD*	*Max*	*Min*
Immediate recall	46	4.24	2.50	10	0
Delayed recall	43	2.04	2.55	10	0

Note. There were 10 items in total; the percentages of recalled contexts were 42.4% in immediate and 20.4% in delayed conditions.

This result suggested that the incidental learning of word-context relation did occur, with a probability of 42.4%. As past studies such as Wesche and Paribakht (1996) assumed, the ability to produce vocabulary with context is an important aspect of mastering vocabulary. The present results provided evidence that the context production in relation to the target items was incidentally achieved through single context exposure. Considering that vocabulary knowledge is gained through repetition of a variety of contextual inputs (Bolger et al., 2008; Webb, 2007b), it is suggested that readers accumulate the memory of contexts they encountered. In addition, the percentage of recalled items in this study was between Webb's (2008) test results of productive and receptive knowledge of meaning, which were respectively 7.2% (the average of 1.3% in uninformative and 13.1% in informative contexts) and 55.8% (the average of 43.8% in uninformative and 67.7% in informative contexts). In his study, receptive knowledge was measured by multiple-choice tests. Considering that the current material contexts were identical to his study, Webb's participants can be assumed to learn around 42.4% of the contexts just as in the present study. If this was the case,

the context memory might well be used in selecting correct choices in a receptive knowledge test. For example, if participants remembered the mental imagery evoked from *I am a doctor at the London Ancon* (e.g., an image of a doctor in a white frock), it should be easy for them to choose the correct meaning of *ancon* from the four choices: (a) *house*, (b) *car*, (c) *city*, and the correct alternative (d) *hospital*. On the assumption that this kind of multiple-choice test measures learners' receptive knowledge of word meanings, the memory of word-context relationship should also be considered as a kind of receptive knowledge.

In terms of the percentage of recalled IUs, the results of the immediate and delayed recall tasks are presented in Table 3.4. The results of recalled IUs generally supported those of the number of contexts reported above. In the immediate condition, participants recalled 29.53% of the whole IUs. From a pedagogical viewpoint, it can be assumed that when EFL learners read short contexts containing unknown words with their translations, and when they encountered these words again, they can recall around 30% of the context information they have read. On the other hand, participants' recall production rate was reduced to 7.98%. Although the percentage itself seems small, it strongly suggests the importance of context reading because the information of context can be retained in learners' long-term memory. Combined with the results of the number of recalled contexts, the data suggested that the word-context connections were rapidly weakened, even though the connections themselves were retained. This result provided the empirical support for the Nation's (2009) suggestion that teachers keep on giving students the contextual input.

Table 3.4

Overall Production Rates of Recalled IUs (%) in Pilot Study 1

Task	*N*	*M*	*SD*	*Max*	*Min*
Immediate recall	46	29.53	19.88	88.33	0.00
Delayed recall	43	7.98	8.73	30.00	0.00

Note. The number of recalled IUs was transformed into a percentage of the total number of idea units in the original passages because the number of idea units differed between passages.

The effects of informativeness and imageability were analyzed using the recall production rates of IUs per context. As mentioned in Section 3.2.4, the material contexts were grouped into the four categories based on their informativeness (informative and uninformative) and imageability (more and less imageable). The overall results from the immediate and delayed recall tasks are presented in Table 3.5 and visualized in Figure 3.1. The figure indicates an interaction between informativeness and imageability, especially in the immediate condition. In the case of the informative contexts, the recall rates were around 30% regardless of their imageability. In contrast, the recall production rates were strongly affected by the imageability factor when the contexts were uninformative. The recall rates in the delayed condition were reduced, but a tendency similar to the immediate recall was found: The imageability effect was larger among the uninformative contexts than the informative contexts.

Table 3.5

Production Rates of Recalled IUs (%) in Terms of the Two Context Factors in Pilot Study 1

			Immediate recall		Delayed recall	
Informativeness	Imageability	*n*	*M*	*SD*	*M*	*SD*
Informative	More	6	29.22	8.51	9.26	7.29
	Less	8	27.85	11.38	4.54	4.78
Uninformative	More	6	41.74	15.64	12.28	10.71
	Less	10	18.30	7.72	0.93	2.27

Note. The production rates were calculated per context.

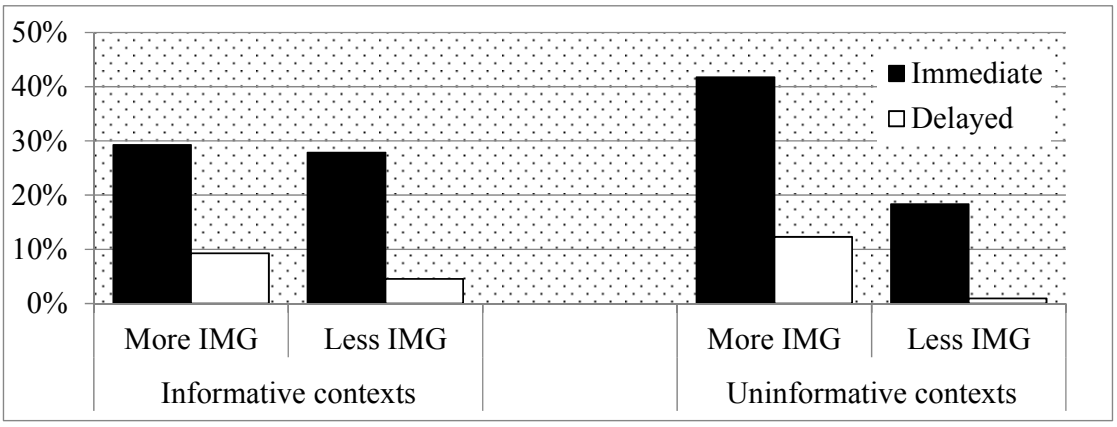

Figure 3.1. Results of the immediate and delayed recall. The learners' recall performance was affected by the interaction of informativeness and imageability. IMG refers to imageability of context.

To verify these observations, a three-way ANOVA was conducted using the recall production rates per context. The independent variables are informativeness (informative and uninformative), imageability (more and less imageable), and the recall timing (immediate and delayed). As a result, the interaction between informativeness and imageability was significant, $F(1, 26) = 5.04$, $p = .033$, $\eta_p^2 = .839$. Post hoc analysis using Bonferroni adjustment indicated that the single main effect of imageability was significant in the

uninformative condition (p = .001) and not significant in the informative condition (p = .515). The single main effect of informativeness was not significant either in the more imageable condition (p = .194) or in the less imageable condition (p = .067). In addition, the main effect of recall timing was significant, $F(1, 26) = 135.51$, $p < .001$, $\eta_p^2 = .839$. This indicated that recall production was significantly reduced by the delay. No other interactions were significant: between informativeness and recall timing, $F(1, 26) = 0.21$, $p = .649$, $\eta_p^2 = .008$; between imageability and recall timing, $F(1, 26) = 1.27$, $p = .269$, $\eta_p^2 = .047$; nor two-way interaction of the three factors, $F(1, 26) = 3.98$, $p = .057$, $\eta_p^2 = .133$. All these results statistically supported the observation of Figure 3.1 mentioned earlier.

The main finding from the ANOVA was the interaction of informativeness and imageability of context. Although Webb (2008) emphasized the importance of text feature, his variable could not explain the recall production rates by itself. The interaction of the two text feature indicates that the function of mental imagery was somewhat different between informative and uninformative conditions. As introduced earlier, mental representation of context can be divided into three levels. Tapiero (2007) argued that when construction of text-based representation is difficult, readers remember the situation model evoked by the text rather than the text itself. In the present experiment, the participants' recall rates were affected by imageability, only when the context informativeness was limited. Whereas the informativeness factor was not directly related to what was called text-based representation, one may argue that the function of imagery factor in the present result seems similar to the situation model. As far as the recall probability of context is concerned, learners might utilize mental imagery when the semantic relationship between context information and the target words was not very strong. However, it was also suggested that the context imageability was the more influential factor than the context informativeness.

3.2.4 Discussion

There were two main findings from the present experiment. The first point was related to the mental connection between target words and contexts, which was represented by the recall rates of contextual information. The results showed the participants incidentally learned the word-context association of 42% of the items. The mental representation of the context might well function as a trigger of multiple-choice alternatives in receptive knowledge tests. However, the connection was weakened by the one-month delay. The second point concerned the two context features: informativeness and imageability. In terms of the mental representation theory, the informativeness factor seemed closer to the semantic or textbase level of representation; the imageability factor was related to a representation called the situation model. The result indicated the interaction of these two factors, suggesting that the learners depended on mental imagery when the presented context had less information. It was consistent with the mental representation theory. Thus, it was shown that learners construct word-context connections in the relation between textual information and their own mental imagery. However, the context imageability was the more influential factor than informativeness. The imageability factor was therefore considered throughout the current research.

Although this experiment was just a pilot study, the above findings provided two pedagogical implications. First, constant contextual input should be promoted, even if each single exposure to context seems to have little effect. Past studies on incidental vocabulary learning reported discouragingly low rates of successful learning. However, the present study revealed that memories of context are retained after reading. Furthermore, such memories of context connected to lexical items can be a part of receptive knowledge of vocabulary meaning, as discussed earlier. Therefore, each encounter with new words in context possibly affects learners' vocabulary knowledge. Secondly, it is advisable for teachers to provide learners with as imageable a context as possible. Hasegawa (2010) suggested that context can

make target words more imageable. In the present data, imageable context never had negative effects on the word-context relationship. Combining these suggestions, incidental vocabulary learning using imageable context should have facilitating effects on the memory of both context and target words. Therefore, reading materials should be imageable if they are used for the purpose of learning new words. With the two major pedagogical implications, this study provides the empirical support for the contextualized learning of vocabulary.

3.3 Experiment 1: Context Imageability Effect

3.3.1 Purpose

Experiment 1 examined how the mental imagery of context affects the intentional learning of unfamiliar words. The imageability data collected in Pilot Study 1 were used to identify more and less imageable contexts. Comparing vocabulary test scores between more and less imageable conditions, the following research questions (RQs) were examined:

RQ1-1: Does context imageability affect vocabulary test scores when learners study target words in a glossed context?

RQ1-2: Does context imageability have different effects on the scores of tests whose contexts are the same as or different from the learning context?

The answer to the first research question is gained by comparison between more and less imageable contexts. This will provide an empirical evidence of whether mental imagery is closely related to contextualized intentional learning of vocabulary. The answer to the second research question is obtained by comparison of the two different test formats. In one of these test formats, the target words were presented in the same contexts as the learning phase. In the other test format, new contexts were used to test the participants' knowledge. The design of the experiment is described in detail in the next section.

3.3.2 Method

3.3.2.1 Participants

A total of 43 university students took part in the experiment. As in Pilot Study 1, their majors were various including engineering, sociology, or international studies. However, data from six students were excluded from analysis, because each of them has experienced of living in an English-speaking country for more than half a year. Thus, data from the remaining 37 EFL learners were used to examine the RQs.

3.3.2.2 Materials

The target words were prepared based on those in Pilot Study 1 (see Appendix 3.4). All 10 words were pseudowords; six of them were nouns and the other four were verbs. In Pilot Study 1, there were 50 contexts in total, where five contexts were prepared for each target word. In Experiment 1, 30 contexts were selected from the 50; 10 of them were more imageable, another 10 were less imageable, and the other 10 were neutral. These three groups of contexts were used as follows. The more and less imageable contexts were used in both learning and testing phases; these learning contexts were then compared in terms of their imageability. On the other hand, the 10 neutral contexts were used only in the testing phase. Using these contexts, it was examined whether contextualized learning effects can be observed in an altered context.

The mean imageability ratings for the three groups of contexts (i.e., more and less imageable contexts and neutral contexts) based on Pilot Study 1 are shown in Table 3.6. The imageability of the three groups of contexts was compared by a one-way ANOVA. Results showed that the intended differences were significant, $F(2, 18) = 45.11$, $p < .001$, $\eta_p^2 = .834$. Post hoc analyses with Bonferroni adjustment showed that the more imageable group had statistically higher ratings than the neutral ($p = .001$) and less imageable groups ($p < .001$), and that the neutral group had statistically higher ratings than the less imageable group (p

= .006). This confirmed the validity of the grouping of the contexts. Using Webb's (2008) contexts, the learning phase of the present study was designed to follow Webb (2007a), which focused on the intentional learning of pseudowords. In Webb (2007a), participants learned 20 target words in eight minutes. To imitate this learning condition, 10 additional sets of words and glossed contexts were selected from Webb (2007a). Thus, the pseudowords used in the present experiment consisted of 10 target words and 10 filler words, which were learned but not analyzed.

Table 3.6

Imageability of the Contexts in Experiment 1

Contexts	Use	*n*	*M*	*SD*
More imageable	Learning and testing phases	10	6.25	0.41
Neutral	Testing phase	10	5.72	0.48
Less imageable	Learning and testing phases	10	4.45	0.75

Note. The imageability ratings were made in Pilot Study 1 based on a 7-point scale.

To investigate the influence of L2 lexical proficiency, the experimenter conducted the vocabulary and expression subsection of the *Eiken Test in Practical English Proficiency* as a proficiency test. This test was adopted because of its practicality and clear target level; for example, Grades 2 and pre-2 is the benchmark level of high school students and Grade pre-1 is the level of university graduates. Results of the proficiency test are briefly reported later.

3.3.2.3 Procedure

Experiment 1 consisted of the learning and testing phases and a lexical proficiency assessment. The participants were given a booklet and asked to perform each task on each page, following the examiner's instructions. The instructions were always made both orally

and visually in Japanese. The instructions for the tasks are presented in detail in Table 3.7. The experimental session started with the general instruction. The overall procedure was completed within approximately 50 minutes. After the experiment, each participant was asked some general questions about their background using a questionnaire.

Table 3.7

Task Instructions in Experiment 1

Task	Instruction
Proficiency test 1	"To complete each item, choose the best alternative from among the four choices."
Intentional learning	"Here is a list of 20 words plus their meanings and example sentences which were made for the purpose of the experiment. Read each context carefully and learn as many words as you can."
New-context test	"To complete each item, choose the best alternative from among the four choices."
Same-context test	(Same as above.)
Proficiency test 2	(Same as above.)

Note. The task instruction was originally given in Japanese.

In the learning phase, students were clearly instructed to memorize the target words and it was announced that there would be several vocabulary tests. This procedure followed the intentional learning paradigm. On the page for this phase, there was a list of 20 words and glossed contexts to be learned. In the list, 10 items were the target items and the other 10 were the filler items to make the situation similar to Webb (2007a). The order of the targets and fillers was randomized. Half of the 10 target words were learned in the more imageable contexts, while the others were in the less imageable contexts. For the purpose of

counterbalancing, two types of booklet (Booklets A and B) were prepared and randomly given to the participants. A target word was presented with a more imageable context in Booklet A when it was presented with a less imageable context in Booklet B and vice versa. Regardless of the booklet types, the participants were asked to learn the items in eight minutes.

In the testing phase, all the students took the two vocabulary tests (hereafter *same*- and *new-context* tests). Both of them were contextualized tests with a multiple-choice format (see Appendix 3.5). The participants were given 10 contexts with a blank and asked to choose the best word for the blank from the four choices. All the choices were pseudowords learned in the former phase. The three distractors for each item always included both noun and verbs, so participants could not choose the correct alternative if they only had knowledge of the part of speech of the target items. Table 3.8 presents examples of the learning and testing context.

Table 3.8

Example of the Combination of Learning and Testing Contexts in Experiment 1

Purpose	Context	Alternatives
Learning context	Perhaps I had a Christmas with my mother once, but I do not denent it. (Target word: *denent* = ～を覚えている)	
Testing context (new)	We liked to talk to him, and we are all very sorry because he is dead. A lot of people are going to (　　) him for a long time.	1. nuggy 2. <u>denent</u> 3. toncop 4. dangy
Testing context (same)	Perhaps I had a Christmas with my mother once, but I do not (　　) it.	1. nuggy 2. <u>denent</u> 3. toncop 4. dangy

The difference between the two tests was in the contexts. In the same context condition, the testing context was identical with the learning context. Therefore, if learners constructed and retained good word-context associations in their minds, they would do well on this test. In other words, the same-context test would measure not only semantic knowledge gain but also the mental associations between the instructed words and contexts. On the other hand, in the new-context condition, the participants were asked to choose the best-matching word for a new context. Therefore, learners' scores on this test would be interpreted as more "flexible" knowledge rather than the memory of word-context associations. Because learners would face many kinds of situation that requires understanding or producing vocabulary in the future, it is necessary for them to learn to adapt their knowledge to another context; such flexibility can be measured by a task with altered contexts (e.g., Anezaki, 2003). To avoid the situation that a learner's memory was reinforced through the re-reading of the learning context, the same-context test was conducted before the new context test. The 10 learning contexts had both five more and five less imageable contexts. Therefore, the contexts used in the same context condition consisted of five more-imageability and five less-imageability contexts. The 10 contexts with neutral imageability were used in the new context test.

3.3.2.4 Scoring

The scores on the vocabulary tests were used in the statistical analyses without any transformation or substitution because all the target items were pseudo words and participants had no prior knowledge of the words. One point was given when a participant correctly chose the target word, and both incorrect choice and making no choice were regarded as an incorrect response.

3.3.2.5 Analysis

Experiment 1 took into account the influence of L2 lexical proficiency In this study, 46

items were adapted from Grades pre-1, 2, and pre-2. The reliability coefficient was Cronbach's α = .73 for the 46 items, and the coefficient became Cronbach's α = .79 after removing relatively 23 less reliable items. According to the sum scores of the remaining 23 items, the participants were divided into upper and lower proficiency groups. Given the mean score of 12.92 and the median of 14, the participants with more than 13 (out of 23) points were regarded as the upper proficiency learners. The numbers of students with average scores are shown in Table 3.9.

Table 3.9

Descriptive Statistics of the Proficiency Test Scores in Experiment 1

Group	*n*	*M*	*SD*	*Max*	*Min*
Upper	19	15.74	1.24	18	14
Lower	18	9.94	3.37	13	0

Note. Maximum possible score = 21.

To examine whether this tendency was statistically reliable, a three-way ANOVA was conducted, using the two test scores. The independent variables were proficiency (upper and lower), imageability of learning contexts (more and less imageable), and testing conditions (same- and new-context). The proficiency factor was a between-subject variable and the imageability and testing conditions were within-subject variables. The alpha level was set at .05.

3.3.3 Results

The upper and lower proficiency learners' scores on the two vocabulary tests (i.e., new- and same-context conditions) are summarized in Table 3.10. Figure 3.2 visualizes the test scores in the same and new conditions, respectively.

Table 3.10

Correct Answer Rates (%) of the Contextualized Recall Tests in Experiment 1

	Same context			New context		
Group	More IMG	Less IMG	Total	More IMG	Less IMG	Total
Upper (n = 19)	80.00	87.37	83.69	70.53	81.05	75.79
	(27.49)	(20.23)	(23.86)	(22.48)	(22.58)	(22.53)
Lower (n = 18)	82.22	68.89	75.56	72.22	67.78	70.00
	(23.65)	(21.93)	(22.79)	(18.33)	(29.22)	(23.78)
Total (N = 37)	81.08	78.38	79.73	71.35	74.59	72.97
	(25.36)	(22.79)	(24.08)	(20.30)	(26.52)	(23.41)

Note. Standard deviations are in parentheses. IMG stands for imageability.

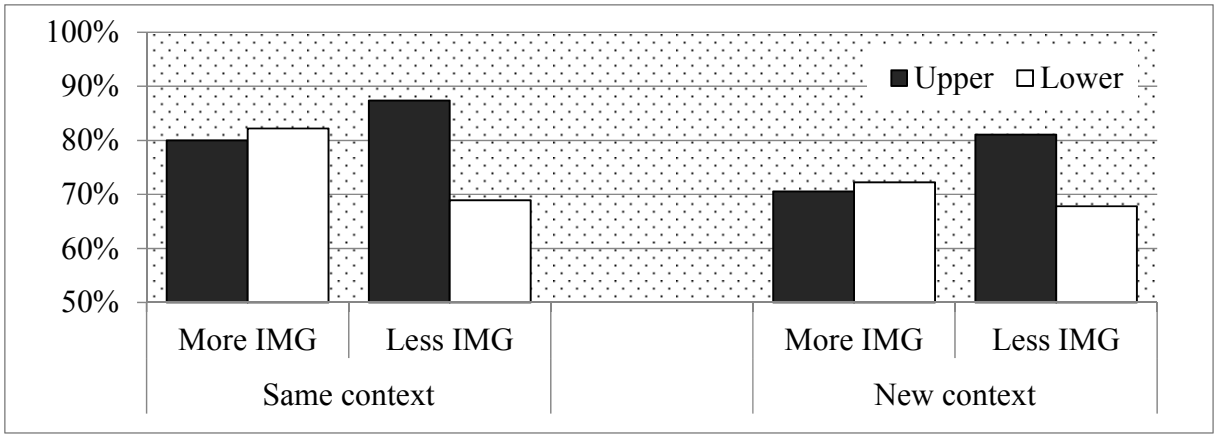

Figure 3.2. Results of the same- and new-context tests.

As shown in Figure 3.2, the results of the two vocabulary tests were similar, in that the lower proficiency group's scores were reduced when the learning contexts were less imageable. To examine whether this tendency was statistically reliable, a three-way ANOVA was conducted, using the two test scores. This ANOVA examined RQs 1-1 and 1-2. First, if the imageability factor significantly affected the test scores, especially those on the same-context test, there would be evidence that learners take advantage of mental imagery from context in an intentional learning task (RQ1-1). In addition, it was important to examine

which proficiency group(s) imageability affects because construction of mental representation during reading a text is an advanced reading process. Secondly, with respect to the testing conditions (RQ1-2), use of the mental associations between learning contexts and target words would be proven when the imageability effects were different between two testing conditions. This RQ was directly related to the interaction between imageability and testing condition factors.

As a result, the interaction between imageability and proficiency was significant, $F(1, 35) = 5.67$, $p = .023$, $\eta_p^2 = .139$. Post hoc analyses with Bonferroni adjustment revealed that the single main effect of proficiency was significant only with the less imageable contexts ($p = .026$, $\eta_p^2 = .134$); it was not significant with the more imageable ones ($p = .773$, $\eta_p^2 = .002$). The single main effect of imageability was not significant among the upper ($p = .096$, $\eta_p^2 = .077$) and lower proficiency learners ($p = .107$, $\eta_p^2 = .073$). This interaction was consistent with the previously mentioned observation of Figure 3.2, where the less proficient learners' scores were apparently lower in the less imageable context. Such a result was assumed to be equal regardless of the testing conditions because the three-way interaction among imageability, proficiency, and testing conditions, $F(1, 35) = 0.27$, $p = .608$, $\eta_p^2 = .008$.

As to the testing conditions, the focused-on interaction with the imageability factor was not significant, $F(1, 35) = 1.19$, $p = .284$, $\eta_p^2 = .033$. No evidence was found that showed that the imageability effect was different according to the testing contexts. On the other hand, the main effect of testing condition was significant, $F(1, 35) = 7.54$, $p = .009$, $\eta_p^2 = .177$. The scores were significantly higher in the same-context condition than the new condition. It was found that learners' performance was better when the word-context relations in the testing phase were identical to the learning phase. This result is not quite surprising, because it was predicted in the *encoding specificity principle*, which argues that memory is most effective when information presented in a learning phase is also available in a testing phase (Tulving & Thomson, 1973).

No other effects were significant: the interaction effect between proficiency and testing condition was not significant, $F(1, 35) = 0.23$, $p = .636$, $\eta_p^2 = .006$; the main effect of imageability was not significant, $F(1, 35) < 0.01$, $p = .994$, $\eta_p^2 < .001$; and the main effect of proficiency was not significant, $F(1, 35) = 1.52$, $p = .226$, $\eta_p^2 = .042$. Note that these insignificant results do not decrease the importance of these two factors. The reason for the insignificance of the imageability and proficiency effects was that these main effects were controlled by the imageability-proficiency interaction.

3.3.4 Discussion

There were two main findings from this experiment corresponding to the two RQs. Both of the RQs were related to the imageability effect on intentional learning of vocabulary in context. Recall that in Pilot Study 1, the context imageability effect was significant both in combination with the informativeness factor and in isolation; it was ensured that the imageability of context affect overall results of context recall after reading. The main difference between Experiment 1 and Pilot Study 1 was in the intentionality of learning. The second experiment here was conducted within the perspective of the intentional vocabulary learning. Therefore, it was examined whether imageability would affect contextualized learning when learners were required to memorize the word meanings.

In the present experiment, imageability effects were controlled by the interaction with the proficiency. There would be two possibilities that the imageability effects were controlled. Firstly, in intentional learning, learners' attention should be paid to the meanings of words. If the Japanese learners of English thought that they were to remember the exact spellings of the given words and the exact translation words, any contextual features would affect learning because their cognitive load was exclusively for the spellings and translations. The actual instruction in the present experiment asked students to read each context carefully and learn as much words as they can. However, there was no evidence that all the participants read all

the contexts very carefully. The other possibility was that learners could not construct good representations of context when they were asked to remember the lexical items. The situation model is supposed to be the highest level representation with the most complicated structure (van Dijk & Kintsch, 1983). These two types of potential difficulty would especially affect learners with lower proficiency. Therefore, an interesting point of view in examining the data was how the context imageability factor interacted with learners' L2 proficiency.

The first research question, or RQ1-1, concerned this issue: Does context imageability affect vocabulary test scores when learners study target words in a glossed context? As a matter of fact, it was indicated that the context imageability influenced the learner's scores according to their L2 lexical proficiency. That is, the context imageability had an interaction effect with the proficiency factor. Furthermore, this study offered evidence that showed context presentation can affect intentional vocabulary learning. Prior research has not provided sufficient evidence that context presentation affects vocabulary learning on condition that L1 translations are also presented (Webb, 2007a). In Webb (2007a), however, many kinds of decontextualized tests were conducted but they never predicted learner performance in a contextualized task. In contrast, the present study conducted two contextualized tests where the word-context associations in their mind would work well. Together with Webb's results, it was found that only contextualized tests can reflect contextualized learning effects when L1 translations are available.

Another point of discussion was the consistency of contexts between the learning and testing phases. The term same-context represents that the identical context was used in both learning and testing phases. On the other hand, a different context was presented in the new-context condition. Whether an interaction between the factor of context imageability and the testing conditions as to context consistency was a key to reveal the difference in results. This issue was related to RQ1-2: Does context imageability have different effects on the scores of tests whose contexts are the same as or different from the learning context? As a

result, it was found that the context imageability effects were not different between the two testing conditions. As argued earlier, if the imageability effect was different between the same- and new-context tests, there would be supporting evidence that learners utilized the mental word–context association intermediated by imagery representation. Recall that in Pilot Study 1, imageable contexts were found to have greater retrieval of the context information. However, the current experiment showed no such interaction effect. What was indicated was the robustness of the context consistency effect. Therefore, it was suggested that the context imageability affects intentional learning but cannot violate the encoding specificity principle.

3.4 Experiment 2: Imagery Instruction Effect

3.4.1 Purpose

Prior research has discovered that the effect of contextualized learning differs according to learner proficiency, but no one has definitively shown under what circumstances lower proficiency learners can effectively gain vocabulary knowledge. Although past studies have suggested that slow learners experience more difficulty than others in achieving contextualized learning, this approach to learning remains important for developing a mental lexicon. Using a contextualized word-recall task as the measure of rate of learning, this study examined the following three factors: learner proficiency, imagery instruction, and context imageability. Focusing on the elaborate processing of context comprehension, the following hypothesis (H) and RQ were addressed:

H2: Japanese EFL learners with higher lexical proficiency are better at recalling unfamiliar words after contextualized vocabulary learning than those with lower proficiency.

RQ2: How are higher and lower proficiency EFL learners' recall of unfamiliar words affected by imagery instruction during learning and context imageability?

The hypothesis will be supported if the recall test scores were higher among the upper proficiency group than the lower proficiency group. This hypothesis will likely be upheld, because the current study is similar to those of Griffin (1992) and Prince (1996). However, it was important to replicate these studies carefully controlling the instruction and context types. To date, although many studies have examined contextualized vocabulary learning, almost no one discussed how the learners' reading process was like. This study aimed to test whether the upper proficiency learners outperform the lower proficiency learners in any situation.

The research question was examined through an analysis of three factors (i.e., Instruction, Proficiency, and Imageability), as well as another factor that was necessary to be taken because of a procedural reason (i.e., Test). If the instruction given the learners to read the context very carefully broadens the lower proficiency learners' attention to context effectively, their recall scores will be higher than in the no-instruction condition. If, on the other hand, the cognitive capacity of the lower proficiency group cannot afford to process so much information, their scores will instead decrease, meaning that imagery instruction is an obstacle to learning. This will imply that the effect of imagery instruction is different according to learner proficiency. Another possible result would be an Instruction-Imageability or Proficiency-Imageability interaction. If these interactions were observed, the conclusion would be that the effect of contextualized vocabulary learning is predominantly controlled by context quality.

3.4.2 Method

3.4.2.1 Participants

A total of 32 undergraduate students in Japan participated in this study. They were majoring in various fields such as engineering, sociology, psychology, and international studies. Most of them were first-year students, but some were in higher years. According to their proficiency test scores, their English skills were estimated to be around the Grade 2 level

or above in the STEP Eiken test. Although there are no formula available to convert STEP Eiken Grade into TOEFL scores, the 32 students seemed to resemble the participants in Prince (1996) in terms of English proficiency, with reference to the official data that students who have obtained STEP Eiken Grade 2 scored 517 points on the TOEIC test on average (Educational Testing Service, 2010). However, data from four participants were excluded from analysis because they knew most of the target words for the main experiment before undergoing the learning phase; the target words were no longer "new words" for them. Thus, data from 28 participants were analyzed.

The participants were divided into two groups according to score on the first subsection of the STEP Eiken test, Grade Pre-1 (2010). Table 3.11 shows the scores on this proficiency test.

Table 3.11

Descriptive Statistics of the Proficiency Test Scores in Experiment 2

Proficiency	*n*	*M*	*SD*	*Min*	*Max*
Upper	13	8.58	2.31	6	13
Lower	15	3.19	1.64	0	5
Total	28	5.50	3.33	0	13

Note. The maximum possible score was 15. The lower proficiency group's minimum score became 0 after excluding the easiest items to improve the test reliability.

This part of the Eiken test was chosen as the proficiency test because it measures lexical skills in the use of vocabulary in context; such skills were considered the most appropriate measure for this study's purposes. The reason Grade Pre-1 was used instead of Grade 2 was that some students got nearly a maximum score when the experimenter presented this easier test in a pilot study. In the current study, the participants answered 25 items in total, but proficiency

scores were calculated based on the 15 items with the highest reliability (Cronbach's α = .81). The 28 students were divided by score into the following two groups: the Upper group, who scored 6 or more (*n* = 13), and the Lower group, who scored 5 or less (*n* = 15).

3.4.2.2 Materials

Given the purpose of the current study, learning materials needed to be at an appropriate level of difficulty. Considering the participants' proficiency level, target words and contexts were developed based on past materials used in STEP Eiken Grade Pre-1 (2010). These learning materials were all different items from the proficiency test items. There were 20 target words in total, including four nouns, eight verbs, and eight adjectives (see Appendix 3.6). All these words were presented to participants in context. In addition, translation words in Japanese that accorded with the context were selected, with reference to an English-Japanese dictionary (Konishi & Minamide, 2001). The contexts had 25.30 words on average (*SD* = 5.74). However, some contexts described concrete situations, but others included relatively little information to imagine any situation; the imageability was different among them. To analyze the imageability difference in contexts, a norming study was conducted as follows.

The norming study was conducted with 10 graduate and undergraduate students who were majoring in applied linguistics or second language acquisition. They were asked to rate the imageability of each context on a 7-point Likert-type scale (7 = *the easiest to imagine a situation described in context*; 1 = *the most difficult to imagine anything about the context*). The context imageability was determined according to the average rating. On this basis, the 20 contexts were divided into More and Less Imageability groups, as in Table 3.12.

Table 3.12

Imageability of the Contexts Used in Experiment 2

Contexts	Use	*n*	*M*	*SD*
More imageable	Learning and testing phases	10	6.22	0.33
Less imageable	Learning and testing phases	10	4.63	0.56
Total		20	5.43	0.93

Note. The possible range of rating was between 1.00 and 7.00. The imageability ratings were made by 10 graduate and undergraduate students who were majoring in applied linguistics or second language acquisition. The instruction given to the raters were the same as in Pilot Study 1. The reliability of the imageability rating was further validated in Pilot Study 2 (see Section 4.2).

3.4.2.3 Procedure

The 28 students each took part in a series of paper-based tasks. The overall procedures are summarized in Figure 3.3. The experiment included a proficiency test, two learning and testing phases, and pre- and posttests. The proficiency test and its results have been described in Section 3.4.2.1. This test was followed by a pretest, where the participants were presented with a list of 20 target words and asked to write their meanings. Almost all these words were unfamiliar to the participants, after the exclusion of four with whom many were familiar (see Section 3.4.2.2). If a participant knew any word in the pretest, the word was excluded from later analyses for that participant.

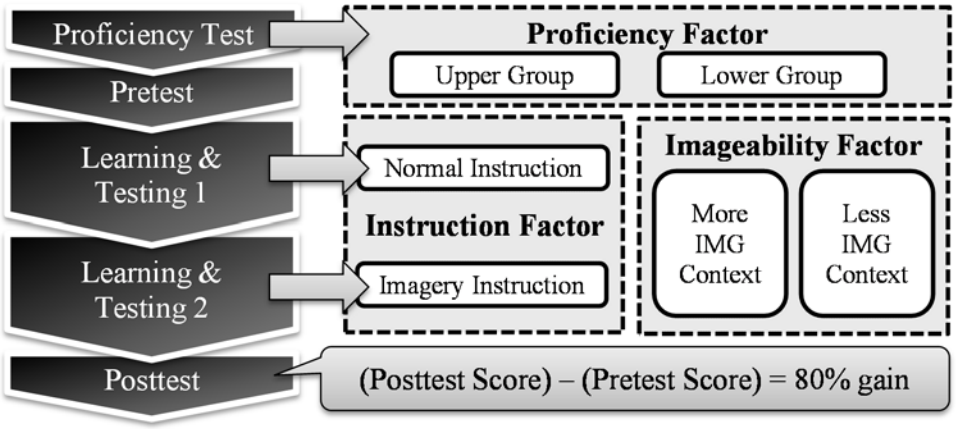

Figure 3.3. The overall procedures and the three main factors in Experiment 2. IMG refers to imageability of context.

Table 3.13

Examples of Learning and Testing Formats in Experiment 2

Phases	Format examples
Learning 1 (Normal instruction)	*bulky* = 大きくて扱いにくい The flight attendant tried to help the woman put her suitcase into the overhead compartment, but it was too bulky to fit.
Learning 2 (Imagery instruction)	*bulky* = 大きくて扱いにくい The flight attendant tried to help the woman put her suitcase into the overhead compartment, but it was too bulky to fit. (least imageable) 1 2 3 4 5 6 7 (most imageable)
Testing 1 & 2 (Contextualized recall)	The flight attendant tried to help the woman put her suitcase into the overhead compartment, but it was too ( ) to fit.
Pre-/Posttests	*bulky* = ( )

As shown in Figure 3.3, the main body of this experiment consisted of two sessions, which included contextualized vocabulary learning under Normal and Imagery instruction

conditions respectively. In the first session, the participants were presented a list of 10 target words in context along with their meanings in Japanese (see Table 3.13). They were then asked to learn the word forms, meanings, and contexts as well as they could within five minutes. Because the tester's intention was fully explained and the participants' attention was on memorizing the list, this task can be regarded as an intentional learning task as described in Hulstijn (2005) and Schmitt (2010). No other special instruction was given in the first learning phase. After learning, the list was removed and a contextualized recall test was conducted in a gap-filling manner (see Appendix 3.7). In this contextualized recall task, the learners needed to remember the word forms matching the contexts, just as in Griffin's (1992) test.

The second learning phase followed the first contextualized recall test. The participants received a new list containing the remaining 10 target words with contexts. To counterbalance any characteristic item effects on the test scores between the two learning conditions, the order of presentation for the two lists was reversed for half of the participants. The procedure was the same as in the first learning phase except for the instruction. As in the first session, the participants' primary task was to learn the word forms, meanings, and contexts within five minutes. However, in the second session, the participants were forced to consider context content and rate the imageability of each context, which confirmed that all participants properly followed the imagery instruction. The format and procedure of the second contextualized recall test was the same as the first. After these two learning/testing sessions, the same test as the pretest was conducted again to ensure that actual vocabulary gain had taken place. Although the scores on this posttest were analyzed together with the contextualized recall tests, the test-type factor was not as important as the other factors.

3.4.2.4 Scoring

In scoring, spelling was not a determining factor if the response could be clearly

understood (see Webb, 2007a). For example, close approximations of the target word *bulky* such as "bulty" and "vulky" were acceptable responses. When two independent raters marked all the responses, the interrater agreement was 97.67%. The disagreements were solved through discussion.

3.4.2.5 Analysis

To examine the effects of imagery instruction and context imageability on vocabulary learning, a four-way ANOVA was conducted using the recall scores of the upper and lower proficiency participants. The independent variables were Proficiency (Upper and Lower), Instruction (Normal and Imagery), Imageability (More and Less), and Test (Contextualized Recall and Posttest). Proficiency was a between-participants factor; the other three were within-participants factors. The alpha level was set at .05 for all analyses; η_p^2 was reported to allow future replication.

3.4.3 Results

The results of the contextualized recall test are presented in Table 3.14 and Figure 3.4; the posttest scores are in Table 3.15. ANOVA results show that the Proficiency × Imageability × Test interaction was significant, $F(1, 26) = 4.47$, $p = .044$, $\eta_p^2 = .147$; it controlled the significant main effect of Test, $F(1, 26) = 17.68$, $p < .001$, $\eta_p^2 = .405$. The interaction suggested that the combination of these factors was quite important and post hoc analyses would be necessary in order to interpret exactly what this interaction indicated.

Table 3.14

Correct Answer Rates (%) of the Contextualized Recall Tests in Experiment 2

Proficiency	*n*	Normal instruction		Imagery instruction	
		More IMG	Less IMG	More IMG	Less IMG
Upper	13	76.28 (28.97)	78.85 (23.64)	67.56 (22.58)	82.05 (18.87)
Lower	15	62.67 (35.96)	55.33 (33.14)	69.56 (34.01)	66.89 (31.03)
Total	28	68.99 (33.03)	66.25 (30.99)	68.63 (28.76)	73.93 (26.21)

Note. The standard deviations are in parentheses. IMG refers to imageability of context.

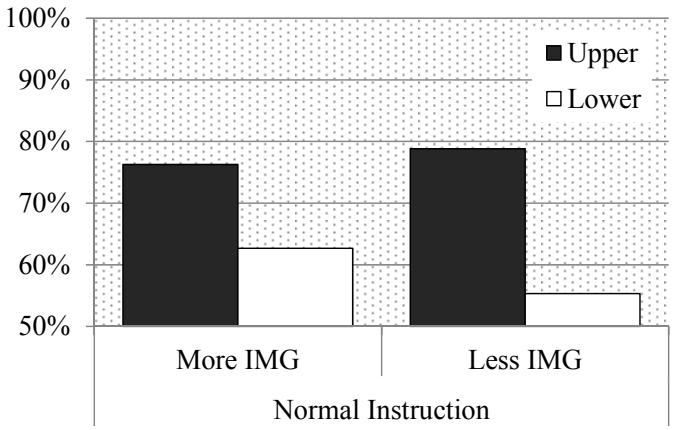

Figure 3.4. Contextualized recall scores (%) as a function of proficiency and imageability in the normal instruction condition. IMG refers to imageability of context.

No other interactions or main effects were significant: (a) a three-way interaction of Proficiency × Instruction × Imageability × Test (p = .878, η_p^2 = .001); two-way interactions of (b) Proficiency × Instruction × Imageability (p = .524, η_p^2 = .016), (c) Proficiency × Instruction × Test (p = .161, η_p^2 = .074), and (d) Instruction × Imageability × Test (p = .069, η_p^2 = .122); one-way interactions of (e) Proficiency × Instruction (p = .577, η_p^2 = .012), (f) Proficiency × Imageability (p = .259, η_p^2 = .049), (g) Proficiency × Test (p = .051, η_p^2 = .139), (h) Instruction × Imageability (p = .985, η_p^2 = .000), (i) Instruction × Test (p = .527, η_p^2

= .016), and (j) Imageability × Test (p = .427, η_p^2 = .024); and main effects of (k) Proficiency (p = .315, η_p^2 = .039), (l) Instruction (p = .706, η_p^2 = .006), and (m) Imageability (p = .892, η_p^2 = .001). Although one of the one-way interactions (g) almost reached the alpha level (i.e., p = .051), it should be noted that it was controlled by the significant two-way interaction of Proficiency × Imageability × Test. The most important finding from these results is that no significant Instruction effects are indicated.

Table 3.15

Correct Answer Rates (%) of the Translation Tests in Experiment 2

		Normal instruction		Imagery instruction	
Proficiency	*n*	More IMG	Less IMG	More IMG	Less IMG
Upper	13	82.05 (18.87)	82.05 (18.87)	84.74 (24.17)	81.67 (18.18)
Lower	15	79.78 (25.81)	85.78 (20.14)	85.33 (23.80)	78.33 (29.74)
Total	28	80.83 (22.47)	84.05 (19.29)	85.06 (23.52)	79.88 (24.67)

Note. The standard deviations are in parentheses. IMG refers to imageability of context.

To interpret the significant interaction (i.e., Proficiency × Imageability × Test), post hoc tests were performed with Bonferroni adjustment. A number of statistical tests were actually done as post hoc analyses, but the meaningful results can be summarized as follows. First, for any factor, significance was found only in the contextualized recall test (p < .05); the posttest scores were equally high, around 80%, which indicates the insignificance of any effects on the test (ps > .10). The posttest results confirmed that the participants in both proficiency groups learned the target items quite well within the limited time provided.

Second, with regard to the contextualized recall scores, a significant simple interaction was found between Proficiency and Imageability; specifically, the effect of Proficiency was significant in the Less Imageability context (p = .031) but not in the More Imageability

context (p = .480), while the effect of Imageability was not significant in either the Upper (p = .129) or Lower (p = .333) learner group. No other effects reached the significance level (ps > .10), except for the Proficiency effect within the Imagery Instruction condition, which showed what is called marginal significance (p = .077). This series of results indicates that the Lower participants' difficulty with contextualized learning was most remarkable in the Less Imageability context.

3.4.4 Discussion

The present study took as points to be examined a hypothesis (H2) and a research question (RQ2). It was hypothesized that the Upper proficiency group would be better at recalling unfamiliar words than the Lower group, based mainly on the results of Griffin (1992) and Prince (1996). The current result supported this assumption, but a further interesting finding was derived: The difference between the Upper and Lower learners was under the control of the Imageability factor, or the ease of raising a mental image of a context. The statistical results indicated that the difference between the two proficiency groups was maximized in contexts whose contents were difficult to imagine. Past studies have indicated that it is difficult for slow learners to learn vocabulary in context; however, the level of difficulty cannot be estimated unless the context quality is carefully considered. The results from the present study strongly suggested that it is essential to examine the ease of evoking mental imagery during learning. This issue is focused on by RQ2 of this study.

With regard to RQ2, this study asked whether and how the Instruction and Imageability factors would affect Upper and Lower groups' scores on the Contextualized Recall test. In addition to the finding described above that Imageability affected the contextualized learning performance, it is now worth discussing how imagery instruction works in contextualized vocabulary learning. Originally, there had been two different predictions about RQ2. The first was that imagery instruction would promote learning, because learners would be able to pay

attention to contextual information under this kind of instruction. This idea was based on the prior suggestion that in a normal learning situation, EFL learners tend to utilize translation instead of context as a source of learning (Webb, 2007a). Because the test format used in this study required participants to make a mental association between a target word and a context, the scores in the Contextualized Recall test would be expected to be higher in the Imagery condition than in the Normal learning situation. The second possibility was that imagery instruction would have a negative influence on word recall. Because vocabulary learners are not very efficient in processing a large amount of textual information (e.g., Laufer & Shmueli, 1997), it could have been troublesome for the participants to be forced to read the context carefully.

In fact, the study found no significant effects concerning the Instruction factor. However, two different interpretations of this finding can be offered. One possibility, of course, is that this kind of instruction has nothing to do with learning efficacy. However, although this was not supported statistically, the data showed that the Lower group's scores increased as much as those of the Upper group only when contexts with good imageability were learned under the instruction to rate the context imageability. This kind of results cannot be explained by the first account. On this point, unfortunately, the present study only found that a marginally significant effect of Proficiency, which was found in the Normal Instruction condition, disappeared in the Imagery condition. This first interpretation might be rejected if future research finds significant difference between Imagery and Normal learning conditions.

The second possibility is that the imagery instruction actually had both positive and negative effects on learning, and that the positive effect offset the negative one in the experiment, which consequently counterbalanced the scores between Imagery and Normal conditions. The combination effects of factors appear to be consistent only with this second interpretation. If the second possibility is true, vocabulary learning under imagery instruction can be regarded as an effective method of learning, because this type of learning ensures that

an advantage can be derived from contextualized learning (Jiang, 2000; see Figure 2.6 in Section 2.4.3) without any decrease in learning efficiency. Further research is needed to examine the exact reason why this kind of effort-requiring instruction did not affect overall test scores.

This study investigated the effects of learner proficiency, teacher instruction, and context quality focusing on contextualized vocabulary learning between upper and lower proficiency groups. The effects of these factors were carefully examined in terms of the mental imagery or semantic representation attached to the items learned, which was assumed to be important in the elaborated learning process. Vocabulary learning in context is supposed to be an important process in the development of a learner's mental lexicon, although it is an effortful task that requires considerable proficiency from the learner and makes considerable cognitive demands.

In pedagogical terms, this study thus sought to discover how a language teacher can reduce students' difficulty in vocabulary learning in context. It supported a rough prediction that Japanese EFL learners with higher lexical proficiency are better at recalling new English words learned in context than those with lower proficiency. However, focusing on context imageability, a surprising result was found: The difference in scores between the two proficiency groups was reduced in highly imageable contexts. Therefore, context imageability should be taken into consideration in future research concerning contextualized vocabulary learning. A pedagogical implication of this result is that teachers should be careful to maintain context quality. By presenting learners with image-evoking contexts, teachers may find that slower learners gain as much vocabulary knowledge as relatively advanced learners.

In Experiment 2, this facilitation effect for the lower proficiency group seemed to be maximized when the imageable context was combined with the instruction to read context carefully and try to imagine the situation described. Further research should clarify the potential effect of learners' effort to develop mental imagery for vocabulary learning contexts.

Applying this research suggestion to an EFL classroom, the current results might be understood as evidence showing that a learner's focus on the meaning of a context does not reduce the efficiency of learning, as far as university students are concerned. Based on the theory that semantic representation from a rich context promotes mental lexicon development, no reason appears to exist to avoid careful reading of a learning context if the learning material is well conceived and of high quality.

3.5 Experiment 3: Contextualized vs. Decontextualized Learning

3.5.1 Purpose

While it is widely argued that vocabulary in a foreign language should be learnt in context, studies have shown that context-based learning is not very efficient when compared to translation-based methods. To examine the effect of a single glossed sentence, Webb (2007a) compared contextualized and decontextualized learning groups, presenting both groups with translations. As a result, there was no significant overall difference between the two groups (i.e., translation-only and translation-and-context groups). The research concluded that a single glossed sentence context may have little effect on vocabulary knowledge (see Section 2.2.1). On the other hand, according to Experiments 1 and 2 in the present study, when learners were provided with both translation and context, vocabulary gain was affected by how well they could imagine the described situation in context. It was necessary to examine if the context imageability effect can be replicated with different materials and participants.

Specifically, there were two weak points in Experiments 1 and 2. First, because these experiments did not compare the more/less imageability condition with the no-context control condition, there was no evidence to argue that context reading is effective for the intentional vocabulary learning. It was necessary to test if a learning condition with more imageable contexts is more effective than a decontextualized condition using a list of only word forms

and translations. Second, the participants in Experiments 1 and 2 were skilled learners of English; they were high-intermediate and advanced learners. Therefore, the effect of reading imageable contexts on vocabulary learning should be further examined with less skilled learners. Additionally, it was unclear whether this imageability effect persists in memory or is merely a short-term tendency; a delayed test of target words seemed highly suggestive. Experiment 3 was designed to explore how context imageability is important in translation-based learning with the following two RQs:

RQ3-1: Do beginner level students learn new words more effectively when they are presented with more imageable contexts than when they are given less imageable, or no contexts?

RQ3-2: Does context imageability have different effects on immediate and delayed test scores?

To test these RQs, Experiment 3 compared three learning conditions: a no-context (i.e., Decontextualized) condition and two contextualized conditions with more and less imageable context, respectively (i.e., the Contextualized+IMG and Contextualized–IMG conditions; IMG stands for imageability, which follows either a plus [more imageable] or a minus [less imageable] sign). As in Experiments 1 and 2, imageability is defined as ease of evocation of mental imagery by written or spoken materials (de Groot, 2011).

3.5.2 Method

3.5.2.1 Participants

For three consecutive weeks, 16 Japanese university students learned 30 low-frequency English words with their Japanese translations across the three conditions. They were regarded as beginner-level learners, but were further divided into a upper-proficiency group (*n*

= 8) and a lower-proficiency group (n = 8) according to their scores on the vocabulary size test (Nation & Beglar, 2007; Japanese version).

3.5.2.2 Materials

A total of 30 words and contexts were prepared for this experiment (see Appendix 3.8). The properties of the learning materials were carefully controlled. The target words were low-frequency nouns and verbs consisting of one or two syllables and word imageability was not biased, in accordance with the MRC psycholinguistic database (Coltheart, 1981). For the contexts, which were based on example sentences from an English-Japanese dictionary (Konishi & Minamide, 2001), eight graduate and undergraduate students majoring in SLA judged whether each context was imageable or not. Imageability measures the quality of a context. Furthermore, according to Pilot Study 1, it also affects how likely students are to remember the contextual meaning. In Pilot Study 1, the eight students rated the ease of imaging for the situation described in each context. Examples of more and less imageable contexts are shown in Table 3.16. The 30 items were divided into three groups for the three learning conditions; for each condition, a word list contained 10 items (i.e., 10 target words paired with translations and example sentences). In addition to the 30 main items, different 30 items were prepared for practice sessions.

Table 3.16

Example Items for Each of the Three Learning Lists Used in Experiment 3

Condition	Target	Translation	Context
Contextualized+IMG	*rite*	儀式 (gishiki)	*The festival comes from an ancient rite.*
Contextualized–IMG	*zeal*	熱意 (netsui)	*They prepared for it with zeal.*
Decontextualized	*proxy*	代理 (dairi)	N.A.

Note. IMG stands for imageability, which follows either a plus (more imageable) or a minus (less imageable) sign.

3.5.2.3 Procedure

The overall procedures are summarized in Figure 3.5. In the experimental session, students took a pretest to confirm that they were unfamiliar with the target words. Next, they were given five minutes to learn three 10-word lists. Each list belonged to one of the following three categories: more imageable context, less imageable context, and no context (see Table 3.16). In this experiment, the learning condition was set as a within-participant factor; all the participants had to learn a list of words in every learning condition. To satisfy this requirement, the participants learned three different lists for three consecutive weeks. On the first week of the main part of this experiment, the participants learned the first word list containing 10 items in the Decontextualized condition. The learning phase was followed by two immediate tests. The first test was the translation test that asked the participants the meanings of the target words (see Appendix 3.9). The second test was the contextualized recall test that asked the participants to fill in a blank in each context that was consistent with those in the learning list (see Appendix 3.10). These two tests corresponded to two tests used in Experiment 5, Decontextualized Test A (recall) and Contextualized Test B (same context), respectively. On the next week, the participants were presented a new list containing different 10 items in the Contextualized+IMG condition. This week, the participants took the

immediate test for the words learned in Contextualized+IMG condition and the delayed test for the words that had been learned a week before in the Decontextualized condition. In this manner, the participants learned all the three lists and took both immediate and delayed test on different weeks.

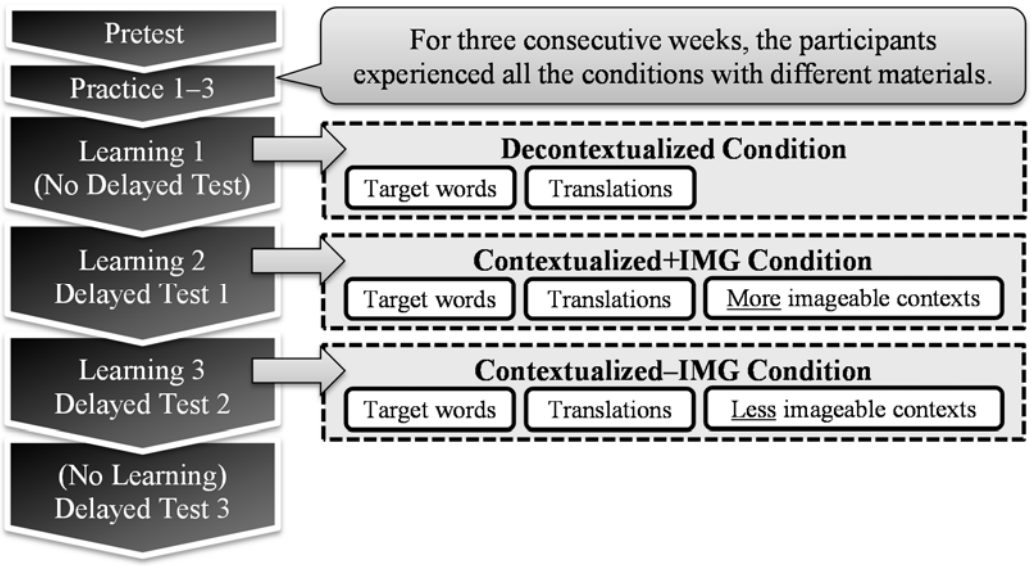

Figure 3.5. The overall procedures and the three learning conditions in Experiment 3. IMG refers to imageability of context.

However, one may argue that this series of procedures was unfair because the participants might become more familiar to the experimental treatments as they experienced learning and testing phases repeatedly. To minimize such an undesirable effect, this experiment provided the participants with a sufficient opportunity of practicing. As illustrated in Figure 3.5, there were three practice sessions in advance of the main learning/testing sessions. Using materials that were not used in any other experimental sessions, the participants learned three different lists with either more/less/no contexts on different weeks. Therefore, it took seven weeks to accomplish Experiment 3 (i.e., Weeks 1–3 for the practice sessions, 4–6 for the main sessions, and 7 for the one-week-delayed test for the Contextualized–IMG condition). The practice session was always accompanied by both

immediate and delayed tests, which ensured that all participants knew about the delayed test by the first day of the main learning session.

3.5.2.4 Scoring

As in Experiments 1 and 2, spelling was not a determining factor as long as the response could be clearly understood. For the translation test, one point was given to an answer that was either the original meaning given in the learning phase or a different expression of the same meaning; however, the participants rarely changed the original wordings that they had learned into new expressions that they came up with.

3.5.2.5 Analysis

Two three-way ANOVAs were conducted using the scores of the translation and contextualized recall tests. For the translation test, the independent variables were Proficiency (Upper and Lower), Learning Condition (Decontextualized, Contextualized+IMG, and Contextualized–IMG), and Test (Immediate and Delayed). On the other hand, the contextualized recall test was not done for the decontextualized learning condition. Although the independent variables were the same as the translation test, the factors were set as follows: Proficiency (Upper and Lower), Learning Condition (Contextualized+IMG, and Contextualized–IMG), and Test (Immediate and Delayed). It should be noted that Experiment 5 had a similar experimental design but it further explored what would happen if learners took a contextualized test after decontextualized learning. Proficiency was a between-participants factor; the other two were within-participants factors. The alpha level was set at .05.

3.5.3 Results

The results of the immediate and delayed translation tests are displayed in Tables 3.17 and 3.18 and visualized in Figures 3.6 and 3.7, respectively.

Table 3.17

Correct Answer Rates (%) of the Immediate Translation Test in Experiment 3

Proficiency	*n*	Decontextualized	Contextualized+IMG	Contextualized–IMG
Upper	8	93.75 (9.16)	85.00 (16.04)	85.00 (14.14)
Lower	8	77.50 (24.35)	65.00 (26.19)	66.25 (25.60)
Total	16	85.63 (19.65)	75.00 (23.38)	75.63 (22.20)

Note. The standard deviations are in parentheses. IMG stands for imageability, which follows either a plus (more imageable) or a minus (less imageable) sign.

Table 3.18

Correct Answer Rates (%) of the Delayed Translation Test in Experiment 3

Proficiency	*n*	Decontextualized	Contextualized+IMG	Contextualized–IMG
Upper	8	18.75 (25.88)	30.00 (19.27)	22.50 (19.09)
Lower	8	3.75 (7.44)	6.25 (5.18)	5.00 (5.35)
Total	16	11.25 (19.96)	18.13 (18.34)	13.75 (16.28)

Note. The standard deviations are in parentheses. IMG stands for imageability, which follows either a plus (more imageable) or a minus (less imageable) sign.

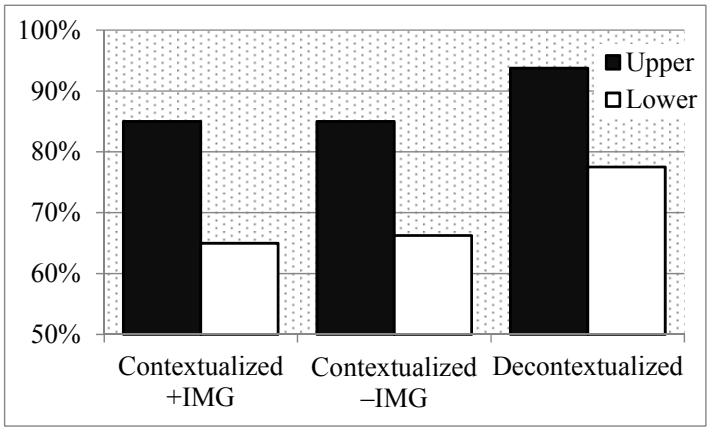

Figure 3.6. The immediate translation results. IMG stands for imageability, which follows either a plus (more imageable) or a minus (less imageable) sign.

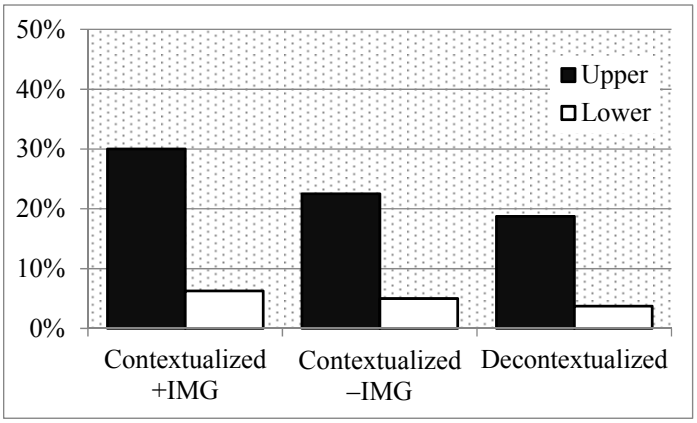

Figure 3.7. The delayed translation results. IMG stands for imageability, which follows either a plus (more imageable) or a minus (less imageable) sign.

Table 3.19

Correct Answer Rates (%) of the Immediate Contextualized Recall Test in Experiment 3

Proficiency	*n*	Contextualized+IMG	Contextualized–IMG
Upper	8	61.25 (26.42)	21.25 (22.95)
Lower	8	25.00 (23.90)	13.75 (20.66)
Total	16	43.13 (30.71)	17.50 (21.45)

Note. The standard deviations are in parentheses. IMG stands for imageability, which follows either a plus (more imageable) or a minus (less imageable) sign.

Table 3.20

Correct Answer Rates (%) of the Delayed Contextualized Recall Test in Experiment 3

Proficiency	*n*	Contextualized+IMG	Contextualized–IMG
Upper	8	2.50 (4.63)	0.00 (0.00)
Lower	8	0.00 (0.00)	0.00 (0.00)
Total	16	1.25 (3.42)	0.00 (0.00)

Note. The standard deviations are in parentheses. IMG stands for imageability, which follows either a plus (more imageable) or a minus (less imageable) sign.

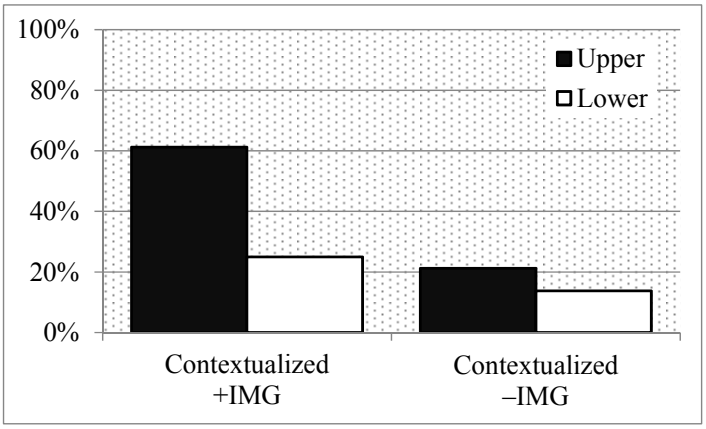

Figure 3.8. The immediate recall results. IMG stands for imageability, which follows either a plus (more imageable) or a minus (less imageable) sign.

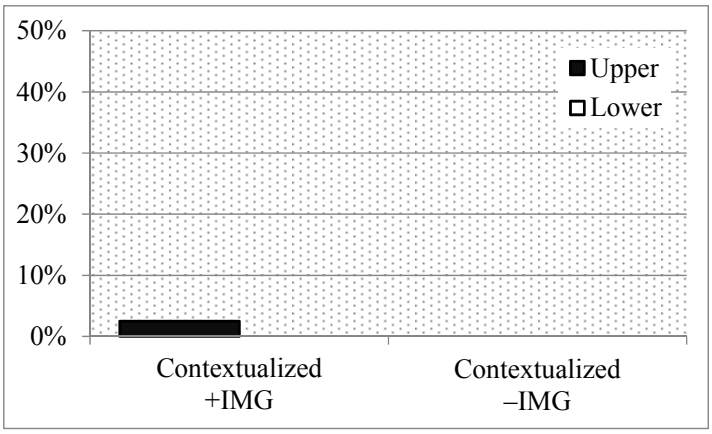

Figure 3.9. The delayed recall results. IMG stands for imageability, which follows either a plus (more imageable) or a minus (less imageable) sign.

The interaction between Learning Condition and Test was significant, $F(2, 28) = 3.93$, $p = .031$, $\eta_p^2 = .219$. This controlled the significant main effect of Test, $F(1, 14) = 247.64$, $p < .001$, $\eta_p^2 = .946$. In addition, the main effect of Proficiency was significant, showing that the upper group outperformed the lower group, $F(1, 14) = 8.86$, $p = .010$, $\eta_p^2 = .388$. To interpret the significant Learning Condition × Test interaction, post hoc tests were performed with Bonferroni adjustment. As a result, the simple main effect of Test was always significant regardless of the learning context conditions (all *p*s < .001). On the other hand, the simple main effect of Learning Condition seemed to be different between the immediate and delayed

tests. First, in the immediate test, Figure 3.6 showed that the Decontextualized condition was better than the other two conditions. Because of this tendency, the difference between Decontextualized and Contextualized+IMG approached the alpha level (p = .059); however, the results were not clear because the sample size was quite small. The differences were not significant between Decontextualized and Contextualized–IMG (p = .271) or between Contextualized+IMG and Contextualized–IMG (p = 1.000). Second, the delayed test showed no significant differences across the learning context conditions: between Decontextualized and Contextualized+IMG (p = .434). Decontextualized and Contextualized–IMG (p = 1.000) or between Contextualized+IMG and Contextualized–IMG (p = .318). The other interactions and main effects were not statistically significant (ps > .05).

With regard to the contextualized recall task, the immediate and delayed test results are displayed in Tables 3.19 and 3.20 and visualized in Figures 3.8 and 3.9, respectively. The two-way interaction of Proficiency × Learning Condition × Test was significant, $F(1, 14) = 7.37$, $p = .017$, $\eta_p^2 = .345$. This controlled the significant interaction of Proficiency × Learning Condition, $F(1, 14) = 10.44$, $p = .006$, $\eta_p^2 = .427$, and Learning Condition × Test, $F(1, 14) = 25.41$, $p < .001$, $\eta_p^2 = .645$, and the main effect of Learning Condition, $F(1, 14) = 30.89$, $p < .001$, $\eta_p^2 = .688$, and Test, $F(1, 14) = 30.98$, $p < .001$, $\eta_p^2 = .689$. However, as shown in Figure 3.9, it was obvious that the delayed test was quite difficult for the learners; no participants in the lower group got any scores on this test regardless of the learning conditions. On the other hand, the more imageable context was more effective for the upper group than the less imageable contexts. Therefore, it was almost meaningless to conduct post hoc analyses for the delayed recall test scores. For the immediate tests, post hoc tests were performed to interpret the significant interaction between Proficiency and Learning Condition. As a result, the imageability effect was found among the upper proficiency group, as follows. First, when the upper proficiency group was focused on, the simple main effect of Learning Condition (i.e., Contextualized+IMG vs. Contextualized–IMG) was significant ($p < .001$).

However, the Learning condition effect was not significant among the lower proficiency group (p = .122). Second, when the Contextualized+IMG condition was focused on, the simple main effect of Proficiency was significant (p = .010). However, the Proficiency effect was not significant for the Contextualized–IMG was not significant (p = .503). Considering both immediate and delayed test scores, the results of the contextualized recall test showed the more imageable context was effective for the upper groups' vocabulary learning compared to the less imageable contexts. The other interactions and main effects were not statistically significant (ps > .05).

In summary, in spite of a few points that were somewhat difficult to interpret, the results of the immediate and delayed posttests showed some differences among the conditions. Namely, while translation-based learning with no context seemed to be more immediately effective than the other two conditions, the contextualized method using imageable contexts showed the greatest effectiveness in the delayed translation test and the immediate and delayed recall tests. In contrast to the results for the imageable contexts, retention in the less imageable context condition was as poor as in the no context condition. Although the differences were not dramatic, the results shown in the four graphs suggest that the effect of context imageability might be larger among the upper proficiency group. The results can be explained in terms of how learners process contexts. When upper-level beginners are provided with imageable contexts, they might be able to process the context information more deeply and consider what is described in the contexts more thoroughly.

3.5.4 Discussion

In this experiment, RQ3-1 asked whether the beginner level students learn new words more effective when they are presented with more imageable contexts than when they are given less imageable contexts and no contexts. The answer to this question might be rather mixed because the results showed different tendencies across test types. Therefore, the answer

to RQ3-1 should be addressed together with the answer to RQ3-2, which focused on the difference of the context imageability effect on the scores of immediate and delayed tests. According to the results of the immediate translation test, the decontextualized learning seemed to be the most effective approach for the upper and lower level beginners. However, in terms of retention of the word meaning and the context-based recall performance, the efficacy of the decontextualized learning was considerably reduced. In particular, the upper group's scores of the recall test were increased when they learned the target words with the more imageable contexts.

The current study suggests that translation is a source of prompt learning. In contrast, the effects of translation-based vocabulary learning are maximized in the longer term, as learners come to better understand how the meaning is realized in context. Given the results from the present study, the term translation-based seems rather ambiguous. Although this type of learning is always based on translations, one may argue that learners do not always experience a mental process of translating. It could be argued that the combination of the upper-proficiency learners and a learning condition that uses imageable contexts makes the most suitable situation for the mental translation in this study. In G. Cook's (2010) words, teachers should train students to use translation as a process rather than a product. For their part, students should compare translations and contexts to understand how word meanings are contextualized.

However, the results of Experiments 1, 2, and 3 should be examined more carefully because Experiment 3 suggested at least two problems to be solved. First, since all experiments in Study 1 used a lexical proficiency test, it was possible that other learner factors such as reading proficiency and learner types in terms of what information learners consider the most during learning (i.e., target words, translations, and contexts). Second, the effect of learner proficiency was not clear because the proficiency effects were observed in a different manner between Experiments 3 and the other two experiments. Because participants

in Experiments 1 and 2 were more advanced than those in Experiment 3, it was necessary to test if the context imageability effect would be found among participants at a medium proficiency level. Additionally, because the sample size was small in Experiment 3, it was meaningful to test if the results showing the context imageability effect can be further replicated in different settings. Study 2 was conducted to deal with all these problems.

3.6 Summary of Study 1

Through a pilot study and three main experiments, Study 1 explored the possibility that learners' mental imagery from context affects intentional vocabulary learning. The H and RQs addressed in Study 1 were as follows:

RQ1-1: Does context imageability affect vocabulary test scores when learners study target words in a glossed context?

RQ1-2: Does context imageability have different effects on the scores of tests whose contexts are the same as or different from the learning context?

H2: Japanese EFL learners with higher lexical proficiency are better at recalling unfamiliar words after contextualized vocabulary learning than those with lower proficiency.

RQ2: How are higher and lower proficiency EFL learners' recall of unfamiliar words affected by imagery instruction during learning and context imageability?

RQ3-1: Do beginner level students learn new words more effectively when they are presented with more imageable contexts than when they are given less imageable, or no contexts?

RQ3-2: Does context imageability have different effects on immediate and delayed test

scores?

These statements include several important factors: (a) context presentation (RQ3-1), (b) context imageability (H2 and all RQs), (c) learner proficiency (H2, RQ2), (d) type of testing context (RQ1-2), (e) test timing (RQ3-2), and (f) imagery instruction (RQ2). Among these factors, the most influential variables were context imageability and learner proficiency. Furthermore, these two factors had an interaction effect on the test scores after contextualized vocabulary learning. The results suggested that participants at a certain level of L2 proficiency were sensitive to the given context type. Based on past studies suggesting the limited effect of contexts in intentional vocabulary learning using translations, one might have predicted that context imageability would not affect participants' test scores at all. However, the current results demonstrated that when university students read the imageable contexts, test scores significantly increased when their proficiency was at a certain level. These students may have built up mental imagery while reading the context and associated it with information about the target word.

However, there were two problems in Study 1. First, since all experiments in Study 1 used a lexical proficiency test, it was possible that other learner factors (e.g., reading proficiency and learner types) related to what information learners considered most important during learning (i.e., target words, translations, and contexts) were given more emphasis than lexical proficiency. Second, the effect of learner proficiency should be further investigated since proficiency effects were observed in a different manner across experiments in Study 1. Specifically, because participants in Experiments 1 and 2 were more advanced than those in Experiment 3, it was necessary to test whether the context imageability effect would be found among participants at a medium proficiency level. These problems were dealt with in Study 2; other remaining problems are discussed in Chapter 6.

Chapter 4

Study 2: Relationship Between Imageability Effect and Learner Proficiency

4.1 Focus of This Study

Study 2 consists of two main experiments and two pilot studies (see Figure 1.1 in Section 1.2). Based on Study 1, L2 learners seemed to be more sensitive to contextual information than expected in previous research. However, it was further suggested that the context imageability effect is closely related to the learners' proficiency in L2. There were two problems in Study 1. First, since all experiments in Study 1 used a lexical proficiency test, it was possible that other learner factors such as reading proficiency and learner types in terms of what information learners consider the most during learning (i.e., target words, translations, and contexts) was more important than the lexical proficiency factor. This problem was tackled in Experiment 4. Second, the effect of learner proficiency should be further investigated because the proficiency effects were observed in a different manner across experiments in Study 1. Specifically, since participants in Experiments 1 and 2 were more advanced than those in Experiment 3, it was necessary to test if the context imageability effect would be found among participants at a medium proficiency level. Therefore, Experiment 5 compared the following three groups of learners: (a) a high-intermediate group, whose proficiency level was estimated to resemble the participants' in Experiments 1 and 2, (b) a low-intermediate group, which was unique in this experiment, and (c) a beginner group, which corresponded to participants in Experiment 3.

As in Study 1, participants in Experiments 4 and 5 learned a list of unfamiliar words in English with translations and example sentences and took some tests including a contextualized recall task of the learned words (see Figure 1.2 in Section 1.3). Through these two main experiments, Study 2 aimed to reveal how the context imageability effect interacts with the learner proficiency factor. Prior to the main experiments, two pilot studies were

conducted to collect data about how learners comprehend contexts. In analysis, Pilot Study 2 focused on the reliability of imageability ratings, whereas Pilot Study 3 examined how well learners can understand the meaning of contexts.

4.2 Pilot Study 2

4.2.1 Purpose

Pilot Study 2 was conducted to validate the materials used in Experiment 2, focusing on the reliability of context imageability judgment. If the imageability ratings were proved to be reliable, Experiment 4 would use the same material set as Experiment 2. In contrast to the materials used in Experiment 1, i.e., pseudowords paired with either informative or uninformative sentence, Experiment 2 used real words and meaningful contexts from an English proficiency test. Since the contexts in Experiment 2 were originally designed for a lexical test for high-intermediate and advanced learners of English, each context seemed to be controlled in terms of how contextual information specifies the meaning of the target word. Therefore, one might argue that the context imageability was not quite different across contexts. To examine the reliability of imageability ratings for the contexts used in Experiment 2, and to prepare the material set for Experiment 4, Pilot Study 2 asked the 10 participants who had rated the context imageability in Experiment 2 to take part in a follow-up session.

4.2.2 Method

4.2.2.1 Participants

As mentioned in Section 3.4.2.2, imageability of the contexts used in Experiment 2 was judged by 10 graduate and undergraduate students in Japan who were majoring in applied linguistics or second language acquisition. The same 10 participants were asked to rate the same contexts again in Pilot Study 2.

4.2.2.2 Materials

The same 20 contexts were used as in Experiment 2 (see Appendix 3.6); they had already been divided into more and less imageable groups. Considering the participants' proficiency level, the materials were developed based on past materials used in STEP Eiken Grade Pre-1 (2010); these were all different items from the proficiency test items. The 20 target words consisted of four nouns, eight verbs, and eight adjectives The target words were glossed using translations in Japanese with reference to an English-Japanese dictionary (Konishi & Minamide, 2001).

4.2.2.3 Procedure

Pilot Study 2 was conducted three weeks after the first session where the participants rated the imageability of each context on a 7-point Likert-type scale (7 = *the easiest to imagine a situation described in context*; 1 = *the most difficult to imagine anything about the context*). To confirm the reliability of the imageability rating, after the order of context presentation was shuffled, the same 10 students were asked to rate the imageability of the same 20 contexts.

4.2.2.4 Scoring

The context imageability was determined according to the average rating. The imageability values were compared between Pilot Study 2 and the norming session in Experiment 2.

4.2.2.5 Analysis

Originally, the 20 contexts were divided into the more and less imageable groups in Experiment 2. To examine reliability of the imageability ratings, two different approaches were applied. First, Pearson's correlation coefficients were calculated for the first and second

sessions. Second, all the contexts were sorted in the order of their imageability in the first and second sessions and the most imageable 10 items were compared between the two sessions.

4.2.3 Results

The imageability ratings in the first and second sessions are summarized in Table 4.1. The ratings in the first and second session were quite similar (r = .91) and the groupings of more and less imageable contexts were exactly replicated. This showed the validity of the imageability ratings in terms of test-retest reliability.

Table 4.1

Imageability of the Contexts Used in Pilot Study 2 and Experiment 4

		First rating		Second rating	
Contexts	*n*	*M*	*SD*	*M*	*SD*
More imageable	10	6.22	0.33	6.16	0.31
Less imageable	10	4.63	0.56	4.61	0.56
Total	20	5.43	0.93	5.38	0.91

Note. The possible range of rating was between 1.00 and 7.00.

4.2.4 Discussion

Based on the results, it was decided that Experiment 4 would use this set of contexts as the experimental materials. Interestingly, although the contexts were originally designed for a testing purpose, not an experimental purpose, the difference in the average imageability ratings between the more and less imageable groups was the largest in the present study; in other words, the difference was larger than the materials used in Experiments 1, 3, and 5. For this reason, the final experiment in the present study (i.e., Experiment 6) used the current material set. Although it was not the main focus of Pilot Study 2, it seemed advisable for

language teachers and learners to know that the contexts presented in a lexical proficiency test are not equally imageable. As mentioned in Pilot Study 1, the context imageability might be independent of the semantic informativeness of contexts (see Section 3.2.4).

4.3 Experiment 4: Lexical and Reading Skills and Learner Types

4.3.1 Purpose

Using the contexts analyzed in Pilot Study 2, Experiment 4 investigates the effect of reading skills and learner types in terms of how learners use a word list during the intentional vocabulary learning. Three experiments in Study 1 suggested that L2 learners can be more sensitive to contextual information than expected in previous research. In addition, it was found that the context imageability effect is closely related to the learners' proficiency in L2. However, since all experiments in Study 1 used a lexical proficiency test to divide participants into groups, it was possible that other learner factors. This experiment aimed to explore this issue.

Study 1 examined how many words learned in context can be recalled in context using two types of contexts: the more imageable contexts that were assumed to evoke mental imagery during reading them and the less imageable contexts that were not as specific, or concrete, as the more imageable ones. In the contextualized recall tests in these experiments, participants had to read and comprehend each context, recall the target word that fits in the context, and then write the word form. While this process, the upper groups in Experiments 1 and 2 were able to match the target words with the appropriate contexts regardless of the context imageability. In other words, they might be able to activate the word-context association in mind that had been constructed in the learning phase. On the other hand, it seemed that the lower groups in Experiments and the beginner-level learners in Experiment 3 could not construct such association in mind during learning and utilize it during taking a test.

Concerning the proficiency effect, there were at least three possible interpretations.

These three possibilities are expressed graphically in Figure 4.1. First, since the upper and lower groups were divided based on their scores on a lexical proficiency test, the most decisive difference between the two groups was in their lexical knowledge and skills. However, compared to the lower group, the upper group might have been more skillful in not only lexical skills but also ability with respect to L2 reading. As mentioned in Section 2.5.1, researchers assume that it is difficult for less skilled learners to utilize contextual information for vocabulary learning because they are not able to read contexts fluently and undergo elaborative processing which requires them to build a mental representation during reading (Griffin, 1992). Therefore, as long as the context imageability effect on test scores is concerned, it was possible that the difference between more and less skillful learners would be clearer when they were divided based on their scores on an L2 reading test.

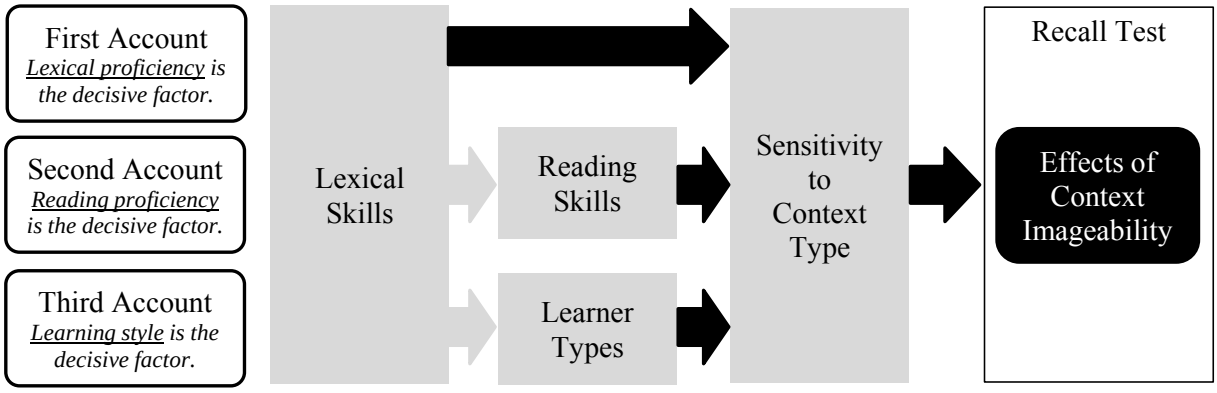

Figure 4.1. Three possibilities of the proficiency effect on the learners' sensitivity to context imageability found in Study 1. The first, second, third accounts regard lexical skills, reading skills, and learner types as the most significant determinants of a learner's sensitivity to context imageability, respectively. Based on the second and third accounts, the participants' lexical skills influenced their sensitivity to context imageability in Experiments 1, 2, and 3 because the learners' lexical skills were mutually related with their reading proficiency and/or learning style.

In addition to these two accounts, there was another possibility that the upper group in Experiments 1 and 2 was different from the lower group not only lexical ability but also their learning styles. In general, more skillful learners can use more vocabulary learning strategies than less proficient learners (see Section 2.4.2; see also Section 6.1.2). In other words, some learners might have concentrated on word forms and translations; others might have read example sentences more carefully. It seemed plausible that the effect of lexical proficiency was observed in Study 1 because the upper and lower proficiency group used a word list differently. Focusing on the high-intermediate and advanced learners in Experiments 1 and 2, it was possible that the upper group's scores were not different across context types because some of them did not read the contexts carefully. If this was the case, the results would show clearer when the participants were grouped according to their learning styles instead of lexical proficiency test scores.

To explore how the learners' sensitivity to the context type relates to the three learner factors, Experiment 4 replicated the intentional vocabulary learning study in a manner that was similar to the normal instruction condition in Experiment 2. To avoid confusion, participants at a lexical proficiency level that was similar to those in Experiment 2 were recruited. The investigation of the low-intermediate and beginner-level learners was carried out in Experiment 5. Two RQs were set as follows:

RQ4-1: Does the effect of context imageability differ between learners with more and less reading skills?

RQ4-2: Does the effect of context imageability differ between (a) word-oriented, (b) context-oriented, and (c) word-and-context oriented learners?

4.3.2 Method

4.3.2.1 Participants

A total of 23 undergraduate students in Japan participated in this study. They were majoring in various fields such as engineering, sociology, and international studies. All of them were first-year students. According to their proficiency test scores, their English skills were estimated to be around the Grade 2 level or above in the STEP Eiken test. The grouping method is described in Section 4.3.2.4.

4.3.2.2 Materials

The materials used in this experiment was the same as in Experiment 2 (see Section 3.4.2.2 and Appendix 3.6; see also Section 4.2). Given the participants' proficiency level, target words and contexts were developed based on past materials used in STEP Eiken Grade Pre-1 (2010); the to-be-learned materials needed to be at an appropriate level of difficulty. As in Experiment 2, these learning materials were all different items from the proficiency test items. Each target word was paired with short verbal context and a translation in the participants' L1 (i.e., Japanese). The imageability of contexts was assessed through a norming study, which is explained in earlier sections; Section 3.4.2.2 describes the first session for norming and Section 4.2 explains the second session for validation.

4.3.2.3 Procedure

The experimental procedures were identical to the normal instruction condition in Experiment 2 (see Section 3.4.2.3), except for an additional proficiency test and a questionnaire on the participant's learning style. The procedures are summarized in Table 4.2.

Table 4.2

The Overall Procedures of Experiment 4

Task	Purpose
Lexical proficiency test	Grouping of the learners (corresponding to the first account in Figure 4.1)
Pretest	Examination of context effect
Intentional vocabulary learning	Examination of context effect
Posttest	Examination of context effect
Questionnaire on learning styles	Grouping of the learners (corresponding to the third account in Figure 4.1)
Reading proficiency test	Grouping of the learners (corresponding to the second account in Figure 4.1)

As the reading proficiency test, Experiment 4 used the second and third subsections of the STEP Eiken test; corresponding to the lexical proficiency test, Grade Pre-1 was adapted for the reading proficiency test. The participants were given four passages; each passage had four questions about the text contents. For both of the lexical and reading proficiency tests, the participants were asked to answer as many questions as possible in 10 minutes.

The questionnaire was given to the participants immediately after the learning and testing, which asked in Japanese, “How did you learn the English words now? Write down your learning method briefly.” However, this question seemed to be vague and learners would be confused about what kind of information they should report. To avoid this, before the main question, a series of questions using a 5-point Likert scale were given to the participants. However, if the experimenter had created the questionnaire items by himself, one may argue that the results of the main question could be distorted by the experimenter’s intention. Therefore, the additional questionnaire was adopted from previous research (Mizumoto &

Takeuchi, 2009). The 25 items about the vocabulary learning strategies were translated into Japanese and presented with the current participants. Examples of Mizumoto and Takeuchi's (2009) questionnaire were "When I try to remember vocabulary, I write it repeatedly," "When I try to remember vocabulary, I say the sample sentence aloud," and "When I try to remember vocabulary, I make a mental picture of what can be associated with a word's meaning" (pp. 448–449). It was expected that after answering to the Likert scale questions, the participants were able to report their own learning styles more objectively (see Appendix 4.1).

4.3.2.4 Scoring

With respect to the contextualized recall test, as in Experiments 1, 2, and 3, spelling was not a determining factor if the response could be clearly understood. For example, when the correct answer was *professing*, spelling errors such as *proffessing* were regarded as acceptable response. In addition, as long as the correct word was answered, it was accepted if the written word was not in the same form as the original; *profess* instead of *professing* was regarded as a correct response.

The lexical and reading proficiency tests consisted of multiple-choice questions. The reliability coefficients showing the two tests' internal consistency were calculated; some test items were excluded because they were found to decrease the test reliability. For the lexical proficiency test, that was the same as the proficiency measure used in Experiment 2, 17 items were used for the main analysis (Cronbach's $\alpha = .86$). For the reading proficiency test, the participants read four passages and solved 16 questions. However, at first, the reliability was not as high as the lexical proficiency test (Cronbach's $\alpha = .59$). A possible reason for this was that the reading test includes more construct than the lexical test. As mentioned in Sections 2.3.1 and 2.4.3, reading comprehension requires many types of knowledge and processes. If Participant A had more background knowledge related to Passage A than others did, her or his would get higher scores for the four items attached to Passage A. Similarly, if Participant B

knew something about the topic of Passage B, she or he could get higher score on this passage than Participant A regardless of their reading fluency. Nevertheless, after excluding four items that were not solved by anyone and four items that were found to reduce the internal consistency of the test, the reliability coefficient of the remaining eight items became much higher (Cronbach's α = .72); these items were used in the main analysis.

4.3.2.5 Analysis

One might argue that the best way of analysis for the current experiment's purpose was multiple regression analysis using the scores of the two proficiency tests. In general, this method looks at the relationship among multiple variables to try to make a prediction about how the independent variables (e.g., the lexical and reading test scores) may predict scores on the dependent variable (e.g., the posttest scores); using multiple regression, one might determine how much of the posttest score variable can be explained by variation in the scores on the different independent variables (see Larson-Hall, 2010). However, since the total number of items were different between the two proficiency tests in the current experiment, it seemed unfair to compare them in their contribution to the posttest scores. Therefore, as in Experiments 1, 2, and 3, the proficiency test scores were used to divide the participants into upper and lower groups.

Experiment 4 conducted three types of analyses corresponding to the first, second, and third approaches mentioned in Section 4.3.1, respectively. First, for the analysis using the lexical proficiency test scores, participants were divided into the Upper and Lower groups based on the average score of 6.83 and the median of 6; eight participants who scored 8 points and above and eight participants who scored 4 points and below were regarded as the upper and lower groups, respectively. With these cut-off lines, the number of participants were equal between the groups.

Second, for the analysis using the reading proficiency test scores, participants were

divided into the Upper and Lower groups based on the average score of 3.26 and the median of 3; nine participants who scored 4 points and above and 10 participants who scored 2 points and below were regarded as the upper and lower groups, respectively. It might appear that the cut-off lines were quite low; however, it should be noted that the most reliable items were used in the analysis after rejecting some unreliable items. With these cut-off lines, the number of participants were almost equal between the groups.

Third, for the analysis based on the learner types, results of the questionnaire on learning styles were examined. However, the results of 25 items using 5-point Likert scales were not considered in the grouping because they were used to familiarize the participants with statements about vocabulary learning styles (see Appendix 4.2). The grouping was based on the participants' responses to the written form question (see Appendix 4.3); three groups shown in Table 4.3 were identified through discussion of two raters. As a result, all participants used the translations representing the word meanings. The most notable difference among the participants was in how much they paid attention to the word forms and contexts. The number of participants in each group was close to the number of participants in the upper and lower groups in the first and second approach based on the lexical and reading proficiency tests.

Table 4.3

The Three Types of Learners in Experiment 4

Group	*n*	Description	Example response
Word-oriented	6	Participants who focused on the word forms (and translations) but did not pay attention to the contexts	"I learned the spellings and meanings. I sometimes tried to pronounce them."
Context-oriented	6	Participants who focused on the contexts (and translations) but did not pay attention to the word forms	"I saw the word meanings and read the example sentences to visualize the image."
Word-and-context-oriented	9	Participants who focused on both word forms and contexts (as well as the translations)	"I repeated looking at the spellings, meanings, and example sentences to learn them."

Note. The example responses were translated into English; all participants write their answers in Japanese. The question did not provide the participants with any alternatives; instead, it asked the participants to report how they had learned the given word list.

Three two-way ANOVAs were conducted using the scores of the contextualized recall tests. For all the three analysis, one of the independent variables was Imageability (More and Less) and the other was concerning learner groups. The first, second, and the third analysis took Lexical Proficiency (Upper and Lower), Reading Proficiency (Upper and Lower), and Learner Type (Word-, Context-, and Word-and-Context) as the learner variables, respectively. The alpha level was initially set at .05; however, considering the small sample size, results were interpreted when the *p* value was below .10.

4.3.3 Results

The overall results are shown in Table 4.4. As in the lower groups in Experiments 1 and

2, the more imagesable context was more effective than the less imageable contexts; the results showed the positive context imageability effect. The results used in the first analysis, which replicated Experiment 2, are displayed in Table 4.5 and Figure 4.2. Also, the results used in the second analysis, which examined RQ4-1 focusing on reading proficiency effects, are summarized in Table 4.6 and Figure 4.3. As shown in these figures, the test scores of the words learned with more imageable contexts were almost the same regardless of lexical/reading proficiency groups. On the other hand, when the words had been learned with the less imageable contexts, the lower proficiency groups' scores became lower than the upper groups; the lower group's recall performance was affected by the context imageability. However, comparing Figures 4.2 and 4.3, it was clear that the lexical proficiency measure predicted the context imageability effect better than the reading proficiency measure, as long as the current results are concerned. Therefore, it seemed that the lexical proficiency was the stronger determinant of whether a learner's scores were dependent on the imageability of contexts.

Table 4.4

Correct Answer Rates (%) of the Contextualized Recall Test in Experiment 4

Context	*N*	*M*	*SD*
More imageable	23	77.83	22.30
Less imageable	23	69.78	27.82
Total	23	73.80	20.90

Table 4.5

Correct Answer Rates (%) of the Contextualized Recall Test as a Function of Lexical Skills in Experiment 4

		More IMG		Less IMG	
Lexical Proficiency	*n*	*M*	*SD*	*M*	*SD*
Upper	8	78.13	22.98	80.63	26.79
Lower	8	76.25	18.47	60.00	26.19

Note. IMG stands for imageability.

Table 4.6

Correct Answer Rates (%) of the Contextualized Recall Test as a Function of Reading Skills in Experiment 4

		More IMG		Less IMG	
Reading Proficiency	*n*	*M*	*SD*	*M*	*SD*
Upper	10	79.00	22.34	74.00	26.75
Lower	9	76.11	26.43	65.00	32.02

Note. IMG stands for imageability.

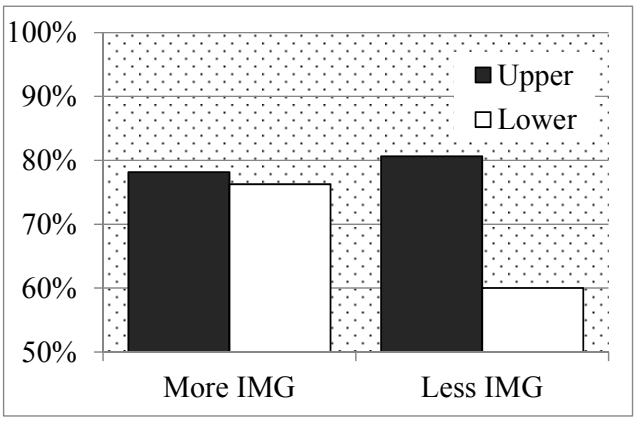

Figure 4.2. The results of the contextualized recall test as a function of the learners' lexical proficiency. IMG refers to imageability of context.

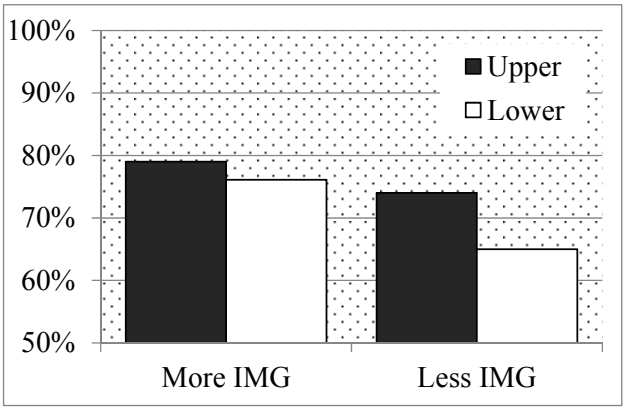

Figure 4.3. The results of the contextualized recall test as a function of the learners' reading proficiency. IMG refers to imageability of context.

However, because of the small sample size, the statistic results were not quite clear. The first ANOVA showed that the interaction between Lexical Proficiency and Imageability was not significant, although it approached the alpha level, $F(1, 14) = 4.11$, $p = .062$. The main effect of Lexical Proficiency was not significant, $F(1, 14) = 1.05$, $p = .323$; the main effect of Imageability was not significant, $F(1, 14) = 2.21$, $p = .159$. Nevertheless, post hoc analyses using Bonferroni adjustment showed that the Lower group's score was significantly different between the More and Less Imageability conditions ($p = .026$), whereas the Higher group's score was not significantly different between the contexts ($p = .708$). This supported the finding from Experiments 1 and 2: Compared to the learners with advanced lexical skill, intermediate learners were more sensitive to the context type during learning a list of target words paired with translation and contexts.

In contrast to the results based on the lexical proficiency scores, the second ANOVA showed that the interaction between Reading Proficiency and Imageability was not significant or approaching the alpha level, $F(1, 17) = 0.19$, $p = .670$. The main effect of Reading Proficiency was not significant, $F(1, 17) = 0.34$, $p = .567$; the main effect of Imageability was not significant, $F(1, 17) = 1.31$, $p = .269$. Taken together, the observation in the first

paragraph of the current section was correct; therefore, the lexical proficiency test in STEP Eiken test was also used in Experiment 5. However, further investigation seemed to be highly necessary with respect to the learner factors because there are many aspects of lexical knowledge such as vocabulary size, depth, fluency. In addition, further examination using other reading proficiency tests should be administered in future research; STEP Eiken Grade Pre-1 might be too difficult for the participants in this study.

The results used in the third analysis, which examined RQ4-2 focusing on learner type effects, are shown in Table 4.7 and Figure 4.4. After all, the third ANOVA indicated no significant interaction between Learner Type and Imageability, $F(2, 18) = 1.45$, $p = .260$. The main effect of Learner Type was not significant, $F(1, 18) = 1.04$, $p = .374$; the main effect of Imageability was not significant, $F(1, 18) = 0.48$, $p = .497$. Therefore, with regard to the three possible accounts of the relationship between learner proficiency and context imageability (see Figure 4.1), the first account seemed more plausible. However, based on the results shown in Figure 4.4, one might argue that the learning strategies during intentional vocabulary learning are related to the results. This issue is further discussed in the next section (see Section 4.3.4).

Table 4.7

Correct Answer Rates (%) of the Contextualized Recall Test as a Function of Learner Types

		More IMG		Less IMG	
Learner type	*n*	*M*	*SD*	*M*	*SD*
Word-oriented	6	63.33	22.51	64.17	24.98
Context-oriented	6	79.17	22.00	83.33	23.38
Word-and-context-oriented	9	81.67	21.21	64.44	31.27

Note. IMG stands for imageability.

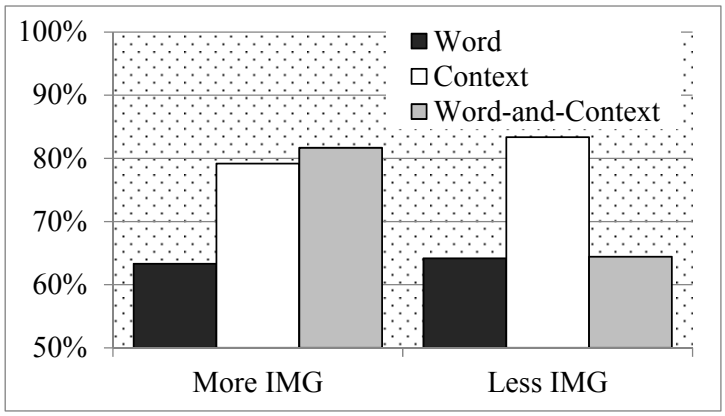

Figure 4.4. The results of the contextualized recall test as a function of the learner types. IMG refers to imageability of context.

4.3.4 Discussion

This experiment examined the effect of reading skills (RQ4-1) and learner types (RQ4-2) on the test scores. When the scores of the lexical proficiency test was taken into account, the results showed that the upper group's scores were consistent regardless of context imageability, whereas the lower group seemed to be more sensitive to the context imageability. When the lexical proficiency factor was replaced with the reading proficiency factor, such a tendency became less obvious. Therefore, as long as the current data are concerned, the second account proposed in Figure 4.1 (i.e., Reading proficiency is the decisive factor) seemed less plausible compared to the first account (i.e., Lexical proficiency is the decisive factor).

However, further investigation should be carried out in future studies because of the following two reasons. First, the present study used the first subsection of the STEP Eiken test, which requires examinees to read short contexts and to choose the best word to fill in a blank; this test requires not only knowledge about the target words but also ability to understand a situation described in context and associate this mental representation to the target word. This process is exactly what this study's intentional learning task required of the participants. It is necessary to test if the aptitude-treatment interaction (ATI; see Section 2.2.1) can be observed

when other tests of vocabulary knowledge are used to assess learners' lexical ability. Additionally, with respect to the reading proficiency test, further replication is necessary because the reading subsections of STEP Eiken Grade Pre-1 might be too difficult for the participants in this study. Nevertheless, Experiment 4 demonstrated that the lexical proficiency test used in Experiments 1, 2, and 3 were the best predictor of the interaction between learner proficiency and context imageability.

On the other hand, the self-report about how the participants used the word list was not a reliable predictor of the interaction of learner factor and context imageability. However, Figure 4.4 seems to suggest that the learning strategies for the word list had some relation with the results to some extent. Although it was not statistically significant, it appears that the word-oriented group learned less regardless of context imageability, and that the context-oriented group learned more regardless of context imageability. Interestingly, the only group whose scores were influenced by the context type was the word-and-context-oriented group. It is worth noting that the word-and-context group was not always the best scorer; although use of all the given information (i.e., word forms, translations and contexts) seems to be an ideal learning habit, this kind of learning might be not effective if the given contexts are not effective. This issue is further discussed in Section 6.2.3, considering the results from other experiment in the current investigation.

One possible theory concerning the unclear effect of learner types in this experiment is that the cognitive strategies during reading each word form, translation, and context were more important than the learner types; this should be explained more because the terms cognitive strategy and learner type should be differentiated. In this experiment, the learner type was determined based on how the participant used the given word list. Specifically, the word-and-context-oriented group tended to learn spellings of the target words accurately and to read the context carefully. However, there are still two possibilities concerning this group's usage of the word list. First, they might read all the target forms and contexts carefully and

repeatedly regardless of the context imageability. Literally, the learners' response such as "I repeated looking at the spellings, meanings, and example sentences to learn them" (see Table 4.3 in Section 4.3.2.5) suggested that they learned all the items in this manner, for no one reported their learning strategies changed according to context type. However, as the second possibility, the word-and-context-oriented learners might read contexts differently between the more and less imageable contexts. If this was the case, a different method should be applied to examine how learners' cognitive strategies differ across more and less imageable contexts. To explore this issue, Experiment 6 conducted a think-aloud study to observe how the word-and-contexts-oriented learners would learn the word forms, translations and contexts.

4.4 Pilot Study 3

4.4.1 Purpose

The primary purpose of Pilot Study 3 was to prepare for materials that were suitable for Experiment 5. As in Pilot Study 1, a group of university students, who did not participated in the main experiment, read a list of short verbal contexts and rated their imageability on a 7-point Likert scale (see Section 3.2.1). Based on their imageability ratings, Pilot Study 3 identified which contexts were more imageable for learners than others. The secondary purpose was to examine whether students at a lower proficiency level, compared to the participants in Experiments 1 and 2, can comprehend the material contexts adequately. In Experiment 5, three groups of learners were distinguished: the high-intermediate, low-intermediate, and beginner groups. Because the main experiment focused on the relationship between learner proficiency and context imageability, it was advisable to test in advance if the learners, especially low-intermediate and beginner-level learners, were able to understand the meaning of the material contexts properly.

4.4.2 Method

4.4.2.1 Participants

Initially, 50 undergraduates took part in this study. They belonged to the same English courses in the same university as the low-intermediate and beginner participants in Experiment 5; because the English classes were assigned according to the students' scores on placement tests and their majors, it was estimated that the proficiency level and background of the current participants were similar to those in Experiment 5 (see Section 4.5.2.1). However, data from 22 students were excluded from analysis because they could not complete the tasks within a limited time. Therefore, data from the remaining 28 students were used in the analyses.

4.4.2.2 Materials

The materials used in this study were (a) 10 target words, (b) 10 more and 10 less imageable contexts for learning, and (c) 10 additional contexts for a contextualized test (see Appendix 4.4). The target words were prepared with reference to Experiment 1, which originally used Webb's (2008) pseudowords. However, for pedagogical reasons, the present study adopted low-frequency words (see Table 4.8), which were not listed in *JACET 8000* (Committee of revising the JACET basic words, 2003), a standard EFL vocabulary list. Japanese translation that accorded with the context were selected from a dictionary (Konishi & Minamide, 2001).

Next, 30 contexts were prepared with reference to the more imageable, less imageable, and neutral contexts in Experiment 1 as well as example sentence in the English-Japanese dictionary (Konishi & Minamide, 2001). Each context was composed of around 10 running words (M = 10.33, SD = 0.71). The contexts used only words in the most frequent 3,000 words in *JACET 8000*, with the exception of the target words. According to calculation by Microsoft Word 2010, the readability measures were as follows: Passive Sentences = 0.0%,

Flesch Reading Ease = 81.4, Flesch-Kincaid Grade Level = 4.4.

Table 4.8

Target Words Paired With Translations Used in Pilot Study 3 and Experiment 5

Target (nouns)	Translation	Synonym	Target (verbs)	Translation	Synonym
esplanade	遊歩道	"street"	*don*	着用する	"wear"
gloaming	夕暮れ時	"evening"	*indite*	執筆する	"write"
repast	食事	"lunch"	*mosey*	ぶらりと訪れる	"visit"
sanatorium	療養所	"hospital"			
tram	路面電車	"train"	*reminisce*	思い出にふける	"remember"
visage	顔つき	"face"			

Note. This study adopted low-frequency words instead of pseudowords used in Webb (2008).

Of the 30 contexts, 10 were selected for greater imageability and another 10 for less imageability. For example, the more imageable context for the target word *reminisce* was *Mike doesn't reminisce about the Christmas day with his girlfriend*; the less imageable context was *Maybe it happened, but she did not reminisce about it*. The additional 10 contexts, which had a middle imageability, were used only in a contextualized test, by replacing a target word with a blank (e.g., *Many people will* _____ *about him for a long time*).

4.4.2.3 Procedure

Pilot Study 3 was not a vocabulary learning experiment; as in Pilot Study 1, the participants were presented with a list of English contexts and asked to read them. The procedure consisted of two tasks: context imageability rating and context translation tasks. The overall procedures are summarized in Table 4.9.

Table 4.9

Summary of the Procedure in Pilot Study 3

Task	Instruction
Context reading	"Read the English contexts carefully."
Imageability rating	"During reading, rate each context as to the ease to arouse an image. If you can rapidly build up a vivid image, circle *7* on the scale. If you cannot evoke any image from a context, choose *1*."
Context translation	"Read the same contexts again. Write the meaning of each sentence in Japanese."

Note. The task instruction was originally given in Japanese.

First, the participants were presented with a list of all the 30 contexts. Each context included one target word, which was glossed with a translation in this session. However, the participants were not informed that the glossed words had any special roles; they were told that these words were glossed because they were rather difficult words. The order of contexts was shuffled so that the participants could read the more, less, and neutral-imageability contexts in a mixed order. As in Pilot Study 1, the participants were asked to rate the words for the ease or difficulty with which they aroused mental images using a 7-point Likert scale. The participants were not informed about the next task until this first task (i.e., imageability rating during reading the list of contexts) was finished. If any preannouncements about the next task had been given to the participants, they would rate each context not for imageability but comprehensibility or ease of translating; this should be avoided in this session.

Second, after the experimenter collected the first lists, a new worksheet was provided to each participant. This new worksheet included the same 30 contexts in the same order as the first list. With this worksheet, the participants wrote the meaning of each context in Japanese. Because the experimenter's interest was in how well the participants comprehend the context

content, the term *meaning*, instead of *translation*, was used in the instruction. Nevertheless, participants regarded this task as a translation task. The two tasks (i.e., the imageability rating and context translation tasks) were completed within 45 minutes.

4.4.2.4 Scoring

In scoring the translation protocols, the present study adopted Ikeno's (1996) criteria, as in Pilot Study 1. All the material contexts were analyzed into IUs by a pair of raters, one was the author and the other was a graduate student who was majoring English education and who was familiar with IU analysis. These raters first independently divided into IUs 10 contexts out of the 30. The interrater reliability was high (percentage agreement = 98.06%). All disagreements were solved through discussion between the raters. Based on the results of this discussion, the remaining 20 contexts were divided by the author.

The participants' recall protocols were scored in terms of the rates of IUs translated correctly. The number of translated IUs was transformed into a percentage of the total number of idea units in the original contexts. To establish valid scoring criteria, the author and a graduate student who was majoring in English education independently scored recall protocols from 10 participants. The interrater reliability was high (percentage agreement = 95.24%). Disagreements were solved through discussion; the remaining data were analyzed by the author.

4.4.2.5 Analysis

The imageability rating data were used to confirm that the more and less imageability contexts worked as intended (i.e., the more and less imageable groups should be regarded as the most and least imageable respectively). However, there were a few cases in which a context regarded as neutral was actually more (or less) imageable than a context regarded as more (or less) imageable for the same target word. In such a case, the context labels (i.e.,

more, less, and neutral) were exchanged across contexts. A *t* test was conducted to compare the mean ratings of the more and less imageable contexts. The alpha level was set at .05.

Concerning the translation data, the average rates were calculated. There were no criteria set for determining how much understanding of textual information was required to reasonably comprehend the overall meaning. However, it seemed helpful to consider Hu and Nation's (2000) frequently cited study, wherein they asserted that "adequate comprehension" requires L2 readers to know 98% of the running words in a text. In their research, an average cued written recall rate of 70.82% was regarded as evidence of adequate comprehension. Based on this assumption, the appropriate expected correct response rate for the context translation results was determined to be approximately 70%.

It is worth mentioning that a similar expectation was set in a recently developed tool for objectively measuring reading comprehension, the *Lexile Framework* (Stenner, 1996; Stenner, Burdick, Sanford, & Burdick, 2006); in this framework, the Rasch model is applied to compare reader ability and text difficulty on a common scale. The framework regards a 75% comprehension rate as the point at which a reader will comprehend enough of a text to understand its complete meaning, a condition Stenner et al. dubbed *targeted reading*. It was established based on Squires, Huitt, and Segars's (1983) observation that reading achievement for second-graders peaked when the success rate reached 75%. However, Squires et al. also maintained that reading comprehension is optimized once approximately 70% of the material has been covered during instruction. Considering these viewpoints, the present study expected the average translation task scores to be around 70% or above.

In addition to analyzing the imageability ratings and context translation protocols, the relationship between these two variables was examined. Specifically, correlation analysis was conducted to test whether imageability judgment was affected by the extent of verbatim understanding of context. If the participants' rating of context imageability was influenced by their comprehension rate, those who earned lower scores in the translation task would not be

able to differentiate between the more and less imageable contexts in the imageability rating task. If this were the case, difference in imageability ratings between the more and less imageable contexts would be minimal among participants with lower comprehension. This seems quite possible in a typical classroom environment, however the context materials were written in plain English, enabling most university students to develop mental imagery from the more imageable contexts with greater ease. To test this, Pearson's correlation coefficient was calculated using (a) the rates of IUs translated correctly, and (b) the differences of imageability ratings between the more and less imageable contexts; each difference was calculated per participant. It was expected that the correlation would be small and/or insignificant.

4.4.3 Results

First, regarding context imageability, the learners' ratings were significantly higher for the more imageable contexts (M = 5.07, SD = 0.62) than for the less imageable contexts (M = 3.56, SD = 0.47), $t(9) = 10.95$, $p < .001$). The ratings of the neutral-imageability contexts were not as high as those for the more imageable items, nor as low as those for the less imageable items (M = 4.38, SD = 0.64). Therefore, it was confirmed that the more, less, and neutral-imageability contexts would work as intended.

Second, based on the results of the translation task, the average rate (%) of IUs translated correctly indicated that the participants had achieved adequate comprehension (M = 70.50, SD = 22.28; see Section 4.4.2.5). The contexts prepared for Experiment 5 were assumed to be comprehensible for L2 learners; however, it was found that some expressions were difficult for the participants to understand. For example, some learners seemed to experience confusion when they read sentences with a center-embedded structure, even if all the running words were quite easy. Therefore, some wordings were revised prior to Experiment 5.

Third, as expected, the difference in imageability ratings between the more and less imageable contexts had no significant correlation with the rates of IUs translated correctly (r = .00, p = .990). In other words, the contexts that were labeled more imageable actually evoked mental imagery with greater ease than the less imageable contexts, regardless of the degree of verbatim understanding of the contexts. In fact, 27 participants out of 28 obtained higher ratings for the more, rather than less imageable contexts. However, participants whose verbatim comprehension rates were higher than others tended to earn noticeably higher ratings for both the more (r = .38, p = .046) and less imageable contexts (r = .37, p = .050, n.s.).

4.4.4 Discussion

The materials for Experiment 5 were validated and improved in Pilot Study 3. First, the more and less imageable contexts were respectively rated as the most and least imageable of the three (i.e., more, less, and neutral) context groups. Second, the participants in Pilot Study 3, who were assumed to be at the same level as the low-intermediate and beginner learners in Experiment 5, seemed to comprehend the contexts at an adequate rate. However, it is worth noting that the context materials used in Pilot Study 3 and Experiment 5 were easier than those in Experiments 2 and 4, which were adapted from the STEP Eiken test.

In addition, the correlation results suggested that participants possessing both higher and lower context comprehension levels were able to distinguish between differences in the ease of evoking mental imagery across context materials. In principal, according to Paivio (1991, p. 256), imageability judgment is affected by the following three factors: (a) the image-evoking value of words and larger language structures, (b) experimental procedures designed to encourage or interfere with the use of imagery, and (c) individual difference in imagery abilities. However, because the current contexts were written in plain English, most participants, including beginners and low-intermediate learners, could develop mental

imagery more easily from the imageable contexts than others.

4.5 Experiment 5: Three Levels of Lexical Proficiency

4.5.1 Purpose

In Study 1, three experiments involving a contextualized test of newly learned words were conducted to examine whether context presentation affects EFL vocabulary learning. The most important finding was that the effect of context types was different between proficiency groups. However, the results seem to be inconsistent among experiments: The context effect was observed in a higher proficiency group in one study but a lower group in another study. Therefore, through a new, larger-scale experiment including both contextualized and decontextualized learning and testing, the present study examined the hypothesis that the most effective way of learning differs by the learner's proficiency level. In particular, since participants in Experiments 1 and 2 were more advanced than those in Experiment 3, it was necessary to test if the context imageability effect would be found among participants at a medium proficiency level. Therefore, Experiment 5 compared the following three groups of learners: high-intermediate, low-intermediate, and a beginner groups.

As in Experiment 3, the current experiment compared contextualized and decontextualized learning by utilizing both contextualized and decontextualized testing. For contextualized learning, two types of context—more and less imageable sentences were prepared, on the basis of the prior studies. The examples of more and less imageable contexts are presented in Section 4.5.2.2. Figure 4.5 illustrates the overall research design, which is described in detail in Section 4.5.2.3.

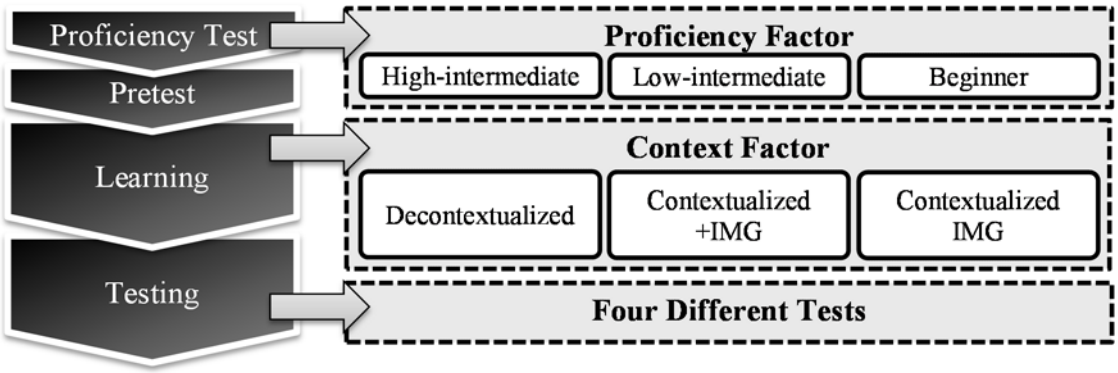

Figure 4.5. The overall procedures of Experiment 5. In the learning phase, three conditions were compared: a no-context (i.e., Decontextualized) condition and two contextualized conditions with more and less imageable context, respectively (i.e., the Contextualized+IMG and Contextualized–IMG conditions; IMG stands for imageability, which was followed by either a plus [more imageable] or minus [less imageable] sign).

Given the results of decontextualized tests reported in Webb (2007a), an interaction between learning condition (i.e., presence or quality of contexts) and learner proficiency was not assumed on the decontextualized tests. Therefore, the following hypothesis was constructed:

H5-1: Scores on the decontextualized tests are affected by learner proficiency, not learning condition.

On the other hand, with regard to contextualized test performance, the following two hypotheses were proposed:

H5-2: Compared to beginners and high-intermediate learners, low-intermediate learners are more sensitive to whether context is given. That is, intermediate learners' scores are higher after contextualized learning than after decontextualized learning.

H5-3: Low-intermediate learners are more sensitive to context quality than beginners or high-intermediate learners. That is, for low-intermediate learners, the more imageable a presented context, the higher their score on a contextualized test.

In summary, the present study expected two types of interaction in the contextualized test results (that is, context-presence × proficiency and context-type × proficiency interactions) and no interaction in the decontextualized test results.

4.5.2 Method

4.5.2.1 Participants

A total of 87 Japanese undergraduates enrolled in EFL classes participated in this study. They were majoring in various physical and social sciences, including international studies, economics, and medical science. Data from three students were excluded from analysis because they happened to skip one or more of the experimental procedures. Thus, the number of participants whose data was retained for analysis was 84.

As shown in Table 4.10, the participants were divided into three groups according to scores on the first subsection of the STEP Eiken test (STEP, 2010; Cronbach's $\alpha = .94$ with 40 items in total), which measures lexical skills on the basis of the use of vocabulary in context. This test was selected as the proficiency measure because skills of this type were considered the most appropriate measure for this study's purposes. Taking into consideration the participants' estimated proficiency levels, Grades Pre-1, Pre-2, and 3 were employed.

Table 4.10

Descriptive Statistics of the Proficiency Test Scores (%) in Experiment 5

		Grade 3		Grade Pre-2		Grade Pre-1		Total	
Proficiency	*n*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
High-intermediate	26	99.74	1.31	99.23	2.72	70.26	12.82	89.74	4.19
Low-intermediate	28	97.62	5.51	93.21	8.19	36.67	12.10	75.83	4.98
Beginner	30	52.67	18.88	34.67	20.63	18.89	11.62	35.41	12.19
Total	84	82.22	25.04	74.17	32.49	40.71	24.40	65.70	24.75

On the test, Grade 3 is a benchmark for junior high school graduates, and is understood to correspond with level A1 (“beginner”) on the Common European Framework of Reference for Languages (CEFR; Council of Europe, 2001); participants’ scores on this grade suggested that the lowest proficiency group in this study could be assumed to be beginners. The highest and middle proficiency groups got almost perfect scores on Grade Pre-2, corresponding with CEFR level A2 (“elementary”), which suggests that these two groups were above the beginner level. However, their scores on Grade Pre-1, corresponding with CEFR level B2 (“upper intermediate”), were noticeably different; thus, the highest and middle proficiency groups were labeled as high- and low-intermediate learners, respectively.

4.5.2.2 Materials

As mentioned earlier (Section 4.4.2.2), 10 target words, 10 more and 10 less imageable contexts for learning, and 10 additional contexts for a contextualized test were used in Experiment 5. The materials were prepared based on Pilot Study 3; the detailed description and examples of the target words, translations, and contexts are provided in Section 4.4.2.2. A total of 30 contexts were prepared with reference to the more imageable, less imageable, and neutral contexts in Experiment 1 as well as example sentence in the English-Japanese

dictionary (Konishi & Minamide, 2001). To maintain comprehensibility of the contexts, all the contexts used only words in the most frequent 3,000 words in *JACET 8000*, with the exception of the target words. Of the 30 contexts, 10 were selected for greater imageability and another 10 for less imageability. Through Pilot Study 3, the ease of evocation of mental imagery of these contexts was rated on a 7-point Likert scale. As a result, their ratings were significantly higher for the more imageable contexts than for the less imageable contexts (see Section 4.4.3). The additional 10 contexts, which had a middle imageability, were used only in a contextualized test, by replacing a target word with a blank.

4.5.2.3 Procedure

The main experiment included a pretest, three types of intentional vocabulary learning task, and two contextualized and two decontextualized tests (see Figure 4.5 in Section 4.5.1). In the pretest, participants were presented with a list of 10 target words and asked to write their meanings. Almost all these words were unfamiliar to the participants; if a participant knew any word in the pretest, that word was excluded from later analyses for that participant.

As shown in Figure 4.5, the following three learning conditions were compared in this study: a no-context (i.e., Decontextualized) condition and two contextualized conditions with more and less imageable context, respectively (i.e., the Contextualized+IMG and Contextualized–IMG conditions; IMG stands for imageability, which follows either a plus [more imageable] or a minus [less imageable] sign). Learners were randomly allotted to one or the other of the conditions. Participants in the Decontextualized condition were presented with a list of 10 target words paired with translations; in the two contextualized learning conditions, participants learned a list consisting of the target words and translations along with contexts. After the tester's intention was fully explained, they learned as many of them as possible within five minutes.

After the learning phase, the list was removed and two contextualized tests (hereafter,

Contextualized Tests A and B) and two decontextualized tests (Decontextualized Tests A and B) were administered (see Table 4.11). The Contextualized Tests were completed by answering gap-filling questions; the participants needed to remember the word forms matching the contexts (Prince, 1996). In Contextualized Test A, all participants were given a common list of 10 new contexts (middle imageability), while in Contextualized Test B, the 10 more imageable contexts were given after Contextualized+IMG learning and the 10 less imageable contexts after Contextualized–IMG learning, in the relevant groups, with each target word replaced with a blank. For the Decontextualized learning condition, the more imageable contexts were presented.

Table 4.11

Examples of Testing Formats in Experiment 5

Test	Cue	Task type	Example
Contextualized Test A	New context	Gap-filling	*Many people will* _____ *about him for a long time.*
Contextualized Test B	Same context	Gap-filling	*Maybe it happened, but she did not* _____ *about it.*
Decontextualized Test A	Target word	Translation	**reminisce** = _____
Decontextualized Test B	Target word	Multiple choice	**reminisce** = (a)… (b)… (c)… (d)…

Note. This is the example for a participant in the Contextuaized–IMG condition; The cue context in Contextualized Test B was the same as in the learning phase.

Decontextualized Test A required the participants to write the correct meaning of each word on a list of target words forms in Japanese. Decontextualized Test B asked them to choose the correct meaning from among four alternatives; the correct answer was the same as the translations used in the learning phase, and the three distractors were translations of other

target words. The number of distractors was determined with reference to past studies (e.g., Sonbul & Schmitt, 2010; Webb, 2008). To avoid learning effects during testing, the four tests were administered in the following order: Contextualized Test B, Decontextualized Test A, Decontextualized Test B, and Contextualized Test A. The test formats were presented in this order in Appendices 4.5, 4.6, 4.7, and 4.8.

4.5.2.4 Scoring

As in Experiments 1 through 4, spelling was not a determining factor in scoring as long as the response could be clearly understood (see Section 4.3.2.4).

4.5.2.5 Analysis

This study had two contextualized and two decontextualized tests. To examine the effects of context and proficiency, two three-way ANOVAs were conducted using (a) the contextualized test scores and (b) the decontextualized test scores. The independent variables were Learning Condition (Decontextualized, Contextualized+IMG, and Contextualized–IMG), Proficiency (Beginner, Low-intermediate, and High-intermediate), and Test (A and B) for each analysis. Condition and Proficiency were between-participants factors; Test was a within-participants factor. The alpha level was initially set at .05.

4.5.3 Results

The test results are summarized in Tables 4.12 to 4.15. ANOVA results showed interesting interactions of factors in the Contextualized and Decontextualized Tests; in particular, it was found that the results of Contextualized Test B should be discussed separately from the other tests. Section 4.5.3.1 explains what was found in Contextualized Test A and the Decontextualized Tests in terms of level of mastery, followed by Section 4.5.3.2 focusing on Contextualized Test B.

The ANOVA results for the Contextualized Tests showed that the Condition × Proficiency × Test interaction was significant, $F(4, 75) = 2.63$, $p = .041$, $\eta_p^2 = .123$; it controlled the significant interaction of Proficiency × Test, $F(2, 75) = 6.32$, $p = .003$, $\eta_p^2 = .144$, and the significant main effects of Proficiency, $F(2, 75) = 46.27$, $p < .001$, $\eta_p^2 = .552$, and Test, $F(1, 75) = 10.05$, $p = .002$, $\eta_p^2 = .118$. No other interactions or effects were significant (*p*s > .10). For the Decontextualized Tests, the interaction for Condition × Proficiency × Test was not significant, $F(4, 75) = 2.05$, $p = .096$, $\eta_p^2 = .099$. Instead, the Proficiency × Test interaction was significant, $F(2, 75) = 14.78$, $p < .001$, $\eta_p^2 = .283$; it controlled the significant main effects of Proficiency, $F(2, 75) = 10.48$, $p < .001$, $\eta_p^2 = .218$, and Test, $F(1, 75) = 94.43$, $p < .001$, $\eta_p^2 = .557$. No other interactions or main effects were significant (*p*s > .10).

Table 4.12

Correct Answer Rates (%) of Contextualized Test A (New Context) in Experiment 5

Proficiency	*n*	Decontextualized	Contextualized+IMG	Contextualized–IMG
High-intermediate	26	86.67 (16.58)	92.09 (9.37)	76.39 (34.17)
Low-intermediate	28	62.50 (38.53)	65.75 (32.58)	61.11 (24.69)
Beginner	30	9.00 (9.94)	21.00 (24.24)	20.00 (21.08)
Total	84	51.16 (40.94)	57.29 (38.02)	51.34 (35.77)

Note. The standard deviations are in parentheses. IMG stands for imageability, which follows either a plus (more imageable) or a minus (less imageable) sign.

Table 4.13

Correct Answer Rates (%) of Contextualized Test B (Same Context) in Experiment 5

Proficiency	*n*	Decontextualized	Contextualized+IMG	Contextualized–IMG
High-intermediate	26	65.83 (25.86)	95.84 (8.26)	55.83 (36.40)
Low-intermediate	28	55.56 (31.32)	51.50 (31.25)	49.44 (29.50)
Beginner	30	22.00 (22.01)	24.00 (21.71)	15.00 (15.09)
Total	84	46.88 (32.01)	54.35 (36.80)	39.20 (32.72)

Note. The standard deviations are in parentheses. IMG stands for imageability, which follows either a plus (more imageable) or a minus (less imageable) sign.

Table 4.14

Correct Answer Rates (%) of Decontextualized Test A (Recall) in Experiment 5

Proficiency	*n*	Decontextualized	Contextualized+IMG	Contextualized–IMG
High-intermediate	26	90.00 (12.25)	100.00 (0.00)	85.56 (18.10)
Low-intermediate	28	76.67 (34.28)	67.00 (21.63)	75.56 (16.67)
Beginner	30	42.00 (30.11)	50.00 (39.72)	73.00 (29.83)
Total	84	68.57 (33.63)	70.36 (33.16)	77.86 (22.50)

Note. The standard deviations are in parentheses. IMG stands for imageability, which follows either a plus (more imageable) or a minus (less imageable) sign.

Table 4.15

Correct Answer Rates (%) of Decontextualized Test B (Recognition) in Experiment 5

Proficiency	*n*	Decontextualized	Contextualized+IMG	Contextualized–IMG
High-intermediate	26	100.00 (0.00)	100.00 (0.00)	94.44 (13.33)
Low- intermediate	28	92.22 (23.33)	96.00 (9.66)	96.67 (7.07)
Beginner	30	85.00 (22.24)	91.00 (15.24)	95.00 (10.80)
Total	84	92.14 (19.12)	95.36 (11.05)	95.36 (10.36)

Note. The standard deviations are in parentheses. IMG stands for imageability, which follows either a plus (more imageable) or a minus (less imageable) sign.

4.5.3.1 Level of Mastery

For Contextualized Test A and Decontextualized Tests A and B, no significant main effects or interactions were found with respect to Learning Condition (*p*s > .10). However, the proficiency effects were different between the tests, suggesting that the level of mastery of the form-meaning link differed across the proficiency groups (see Figure 4.6).

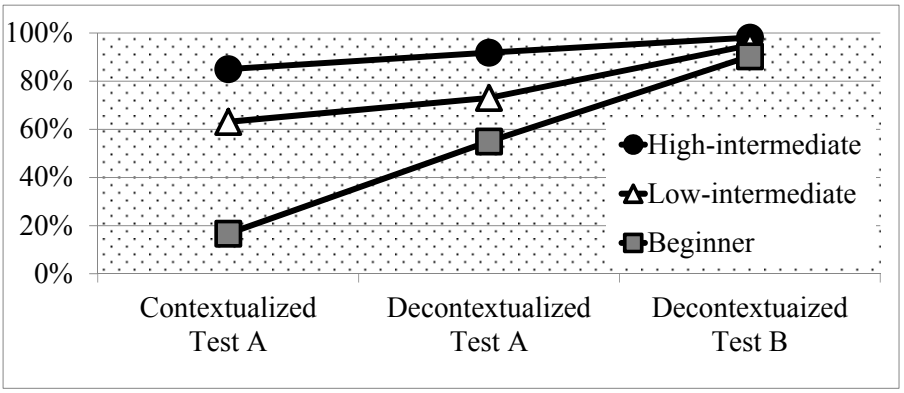

Figure 4.6. Correct answer rates (%) of Contextualized Test A and Decontextualized Tests A and B as a function of proficiency.

For Contextualized Test A, the simple main effect of Proficiency was significant (*p*

< .001). When the three proficiency groups were compared with Scheffe's method, the High-intermediate group's scores were significantly higher than those of the Low-intermediate (p = .011) and Beginner groups (p < .001), and the Low-intermediate group outperformed the Beginner group (p < .001). For the Decontextualized Tests, the differences on Test A were significant between Beginner and Low-intermediate (p = .029), Beginner and High-intermediate (p < .001), and Low- and High-intermediate (p = .028); Test B showed no significant differences between groups (ps > .10).

When the three proficiency groups were separately analyzed, the effect of Test was significant among the Beginner and Low-intermediate groups (both ps < .001) but not the High-intermediate group (p = .112). This indicates that all the learners gained knowledge of the form-meaning link at the initial stage of mastery through intentional vocabulary learning, but that the level of mastery was different across proficiency groups. The ANOVA results for the Contextualized Tests showed that the Condition × Proficiency × Test interaction was significant, $F(4, 75) = 2.63$, $p = .041$, $\eta_p^2 = .123$; it controlled the significant interaction of Proficiency × Test, $F(2, 75) = 6.32$, $p = .003$, $\eta_p^2 = .144$, and the significant main effects of Proficiency, $F(2, 75) = 46.27$, $p < .001$, $\eta_p^2 = .552$, and Test, $F(1, 75) = 10.05$, $p = .002$, $\eta_p^2 = .118$. No other interactions or effects were significant (ps > .10).

In summary, Contextualized Test A and Decontextualized Test A differentiated all three learner groups (Beginner, Low-intermediate, and High-intermediate), whereas Decontextualized Test B shows a ceiling effect. In terms of the level of mastery of the form-meaning link (Laufer et al., 2004; Schmitt, 2010), only High-intermediate learners were able to gain this kind of knowledge at the final level. It was too difficult for the beginners to apply their knowledge of newly learned words to a new context, but that as they improve their proficiency level, they will become gradually able to acquire vocabulary knowledge at the highest level of mastery. Furthermore, because learning condition did not affect the learners' performance on the decontextualized tasks, the data fully support H5-1: Scores on the

Decontextualized Tests are affected not by learning condition but by learner proficiency.

4.5.3.2 Aptitude Treatment Interaction

Scores on Contextualized Test B are displayed in Figure 4.7. ANOVA results showed that the Condition × Proficiency interaction closely approached the alpha level (p = .056), which would indicate that it controls the significant main effect of Proficiency (p < .001). When the results for the three proficiency groups were analyzed separately with Bonferroni adjustment, a significant simple main effect of Condition was found only in the High-intermediate group; Contextualized+IMG was significantly better than Contextualized–IMG (p = .007) but the difference between Contextualized+IMG and Decontextualized was only marginally significant (p = .060), and the difference between Contextualized–IMG and Decontextualized was not significant (p = 1.000). In contrast, the simple main effects of Condition were not significant in the Beginner and Low-intermediate groups (both ps = 1.000).

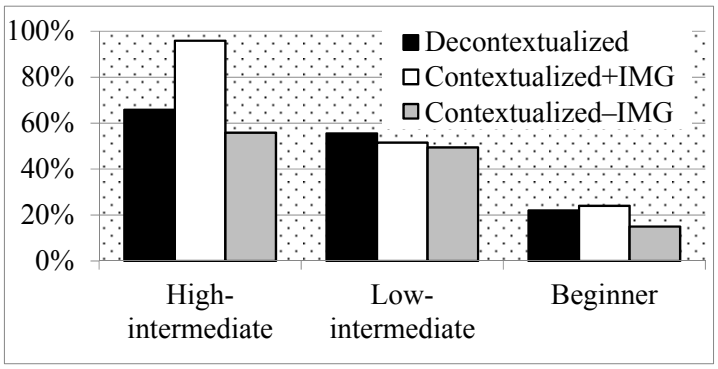

Figure 4.7. Correct answer rates (%) of Contextualized Test B as a function of Learning Condition and Proficiency.

This result can be interpreted as showing that the positive effect of presenting good contexts during translation-based intentional learning of vocabulary was observed only when

the learners were above a certain level of lexical proficiency. Thus, it supports Prince (1996) and Griffin's (1992) contention that contextualized vocabulary learning is more suitable for upper-level learners, and further, suggests that there might be a kind of threshold between the Low- and High-intermediate groups in term of the learners' sensitivity to context.

This observation was further verified when the differences between proficiency groups were examined across learning conditions. In the Decontextualized learning condition, the Low- and High-intermediate groups' scores were significantly higher than those of the Beginner group (*p*s = .019 and .001, respectively), but no significant difference was found between the Low- and High-intermediate learners (p = 1.000). Similarly, within the Contextualized–IMG condition, the Low- and High-intermediate groups were significantly better than the Beginner group (*p*s = .015 and .003, respectively) but there was no significant difference between the Low- and High-intermediate groups (p = 1.000). In contrast, within the Contextualized+IMG condition, the High-intermediate groups significantly outperformed the Low-intermediate and Beginner groups ($p < .001$ and $p = .002$, respectively); the difference between the Low-intermediate and Beginner groups became less clear ($p = .062$). For the Decontextualized Tests, the interaction for Condition × Proficiency × Test was not significant, $F(4, 75) = 2.05$, $p = .096$, $\eta_p^2 = .099$. Instead, the Proficiency × Test interaction was significant, $F(2, 75) = 14.78$, $p < .001$, $\eta_p^2 = .283$; it controlled the significant main effects of Proficiency, $F(2, 75) = 10.48$, $p < .001$, $\eta_p^2 = .218$, and Test, $F(1, 75) = 94.43$, $p < .001$, $\eta_p^2 = .557$. No other interactions or main effects were significant (*p*s > .10).

These results did not support H5-2 or H5-3; however, the ATI was observed in a different manner. That is, the statements should be modified as follows: Compared to the beginners and <u>low</u>-intermediate learners, <u>high</u>-intermediate learners are more sensitive to the presence and quality of context. The results from Contextualized Test B clarify the reason for the confusing results among the prior studies; an effect of context quality was consistently found in Experiments 1 and 2 because the participants in those studies fell into the range of

high-intermediate learners. When learners at this proficiency level are further divided into upper and lower groups, the effect of learning condition (specifically, context imageability) is more noticeable in the lower group, because the upper-high-intermediate, or more succinctly, the advanced learners were able to learn the words on a given list regardless of context type. On the other hand, in Experiment 3, a few students in the upper group got almost full scores on the beginner-level proficiency test. Those student might have performed similarly to the high-intermediate learners in the current study, affecting the overall results of Experiment 3, which was a small-scale experiment ($N = 17$).

4.5.4 Discussion

The present study found an ATI between the contextualized learning of vocabulary and learner proficiency, in an intentional, sentence-based, and translation-given learning situation. The results suggest that even in a translation-based learning condition, contextual information is important for high-intermediate EFL learners. In contrast, beginners and low-intermediate learners might depend on a given translation regardless of the given context. The results also indicated that these groups' gain in the form-meaning knowledge was at the halfway level of mastery. The context effect might also disappear among the advanced learners if they are able to learn the target words almost perfectly. In Prince (1996), the upper and lower proficiency groups' TOEFL scores were 480 (equivalent to CEFR level B2) and 397 (level B1), respectively; these results are quite consistent with the assumption mentioned above if we regard Prince's upper and lower groups to be equivalent to high- and low-intermediate learners. However, the proficiency effects found in the current research are further discussed in Sections 6.1.2 and 6.2.1. These sections summarize the results and reexamine how learner proficiency relates to the imageability effect on the intentional vocabulary leaning.

In pedagogical terms, it will be informative for teachers' practice to be aware that learners at different proficiency levels have different degrees of sensitivity to context. When

students have developed their lexical proficiency to a certain level, they might be able to compare a new word's meaning with the situation described in context. In such a case, their vocabulary knowledge gain will be maximized when a good context is provided. The current results indicate that a short sentence made up of around 10 frequent words and conveying richly specific and meaningful information may facilitate the students' future performance on a contextualized task. On the other hand, decontextualized learning might be an efficient approach for beginners who are not able to deal with contextual information during intentional vocabulary learning (Griffin, 1992; Prince, 1996). However, of course, intentional, translation-based learning is not enough for adequate vocabulary development. Future research should examine whether the aptitude-treatment interaction found in the present study can be applied to other approaches to vocabulary learning.

4.6 Summary of Study 2

Through two pilot studies and main experiments, Study 2 examined two problems addressed by Study 1. First, since all experiments in Study 1 used a lexical proficiency test, it was possible that other learner factors (e.g., reading proficiency and learner types) related to what information learners considered most important during learning were given greater emphasis than lexical proficiency. Second, because the proficiency effects were observed in a different manner across experiments in Study 1, it was necessary to test if the context imageability effect would be found among participants at a low-intermediate proficiency level. To examine these issues, the following RQs and Hs were set in Study 2:

RQ4-1: Does the effect of context imageability differ between learners with more and less reading skills?

RQ4-2: Does the effect of context imageability differ between (a) word-oriented, (b) context-oriented, and (c) word-and-context oriented learners?

H5-1: Scores on the decontextualized tests are affected by learner proficiency, not learning condition.

H5-2: Compared to beginners and high-intermediate learners, low-intermediate learners are more sensitive to whether context is given. That is, intermediate learners' scores are higher after contextualized learning than after decontextualized learning.

H5-3: Low-intermediate learners are more sensitive to context quality than beginners or high-intermediate learners. That is, for low-intermediate learners, the more imageable a presented context, the higher their score on a contextualized test.

Compared to Study 1, whose RQs concerned six factors in total, only a few variables were considered in Study 2. In particular, Study 2 focused on the ATI between the context imageability and the learner groups. However, the methods of grouping leaners were more carefully examined in Study 2.

The interaction effects of the context imageability and learner groups on the test scores were most clearly observed when participants were grouped according to the lexical proficiency test scores. Although more elaborate replication was necessary before deciding which kind of L2 skill was most influential (see Sections 6.2 and 7.2), it became obvious that the lexical proficiency subsection of the STEP Eiken test should be used throughout the present investigation (corresponding to RQs 4-1 and 4-2).

Using this measurement, the ATI between context imageability and learner proficiency was investigated among (a) beginner, (b) low-intermediate, and (c) high-intermediate level learners. Although the exact pattern of the results did not match the author's expectations, they indicated that even in a translation-based learning condition, contextual information is important for high-intermediate EFL learners. Taking these results, as well as Experiments 1 to 3 into consideration, it seems reasonable to assume that there are three stages of lexical

development wherein (a) a learner is unable to process contextual information during intentional vocabulary learning (beginner and low-intermediate levels), (b) the individual becomes highly sensitive to context quality (high-intermediate level), and (c) the learner is able to perpetually learn new words, regardless of context type (advanced level).

Interestingly, this difference between learner groups seems to be independent of the degree of context comprehensibility. One might surmise that the beginner and low-intermediate learners were not as sensitive to context type because they were unable to differentiate between which contexts were more imageable than others, since they could not comprehend the meaning of the given contexts. However, this argument is implausible as far as the current experiment is concerned. According to the supplemental analysis conducted in Pilot Study 3, participants at both higher and lower levels of context comprehension were able to distinguish between differences in the ease of evoking mental imagery across context materials (see Sections 4.4.3 and 4.4.4). In other words, the current material set was properly controlled in terms of the ease of verbatim understanding of the contexts.

As suggested in Section 2.4.1, more imageable contexts are likely to be represented in a reader's mind as an action picture, where the meaning of the entire sentence forms an organized unit (i.e., mental imagery). In addition, the cognitive processes of reading might differ between more and less imageable contexts. According to some traditional studies, when learners read an imageable sentence, they tend to retain the overall meaning; on the other hand, when they read a less imageable sentence, they tend to remember the superficial wordings (Begg & Paivio, 1969; Rowe et al., 1978). However, Studies 1 and 2 did not provide evidence concerning which of the high-intermediate learners' intentional vocabulary learning mechanisms were affected by context imageability. Therefore, to explore this issue, Study 3 further examined how the learning processes were different between the more and less imageable conditions; other remaining problems are discussed in Chapter 6.

Chapter 5

Study 3: Cognitive Strategies of Context Reading During Vocabulary Learning

5.1 Focus of This Study

5.1.1 Overview of Study 3

Study 3 consists of a main experiment and a pilot study (see Figure 1.1 in Section 1.2). Experiment 6 aimed to test a few hypotheses as to how the context imageability factor affects cognitive processes involved in vocabulary learning. To examine the cognitive strategies during reading the more and less imageable contexts for the purpose of vocabulary learning, Study 3 used the think-aloud method that asked participants to report everything they came up with during learning. As introduced in Section 2.5.1, researchers suppose that think-aloud protocols provide a direct view of a reader's mental activity and regard this method as one of the most helpful approach to study cognitive strategies. Based on past studies using this method (e.g., de Bot, Paribakht, & Wesche, 1997; Horiba, 2013; Kaakinen & Hyönä, 2005; Nassaji, 2006; Rott, 2005), the present study assumed that the think-aloud protocols reflect learners' word and context processing strategies, or the *quality of processing* (Rott, 2005, p. 98). Because think-aloud data are most informative about the mental process when learners have difficulty in understanding what they are reading (Block, 1986), it seemed effective to utilize this method in the study of intentional learning of unfamiliar words in L2. This study uses the term cognitive to cover a wide range, referring to learner's actual performance involved in reading and learning the word forms, meanings, and contexts (see V. Cook, 1997, p. 279). Pilot Study 4 was conducted to determine how to analyze and categorize the collected think-aloud protocols. In addition, to discuss the findings from Study 3 more carefully, supplementary data from an additional participant are briefly described in the general discussion (Section 6.1.3).

5.1.2 Background

Studies 1 and 2 found that the L2 learners might become more sensitive to the contexts given in a vocabulary learning task as they become more proficient in the target language. Before describing the method and results of Experiment 6, it should be necessary to review the focus of this thesis. The present research investigates the intentional vocabulary learning because this approach is one major source of lexical knowledge for L2 learners. Among many variations of the intentional vocabulary learning, the most typical type might be the rote memorization of a list of word forms in L2 and their translation into the learner's L1 (see Section 2.2.1), which is often accompanied by a preannouncement of a subsequent test on how well they learned the target words (Hulstijn, 2005; Schmitt, 2010). This type of learning is also often referred to as *paired-associate learning* (e.g., de Groot & Keijzer, 2000; Steinel et al., 2007), for the learners' task is to learn one-to-one associations of L1 and L2 translation equivalents. Researchers have revealed the relative efficacy of intentional vocabulary learning using translations compared to other methods such as lexical inferencing from context. Therefore, intentional vocabulary learning is now regarded as an effective means of lexical development as long as learners also have sufficient opportunities to encounter the new words repeatedly in other contexts (see Nation, 2013).

In addition to the translation pairs, such a word list for intentional vocabulary learning often includes short verbal contexts written in L2. Although a single context can only comprise a limited amount of input, it seems quite reasonable to supply contexts for intentional vocabulary learning because the meanings of various words are conveyed and realized through context. For example, students might learn English word forms such as *apprentice* with Japanese translations such as 見習い [minarai] as well as short verbal contexts such as *Before John became a professional chef, he spent many years working as an <u>apprentice</u> to a master chef at a famous restaurant in London* (see Section 5.3.2.2). However, the effects of glossed contexts have rarely been examined; it is unclear whether and how the

mental process of context reading is necessary during intentional vocabulary learning. As far as the present author is aware, no past studies have clarified whether and how the process of reading a context influences learners' lexical development. In fact, when Webb (2007a) compared translation-based learning with and without a context, the learners' test scores showed no overall difference between these two conditions (i.e., contextualized vs. decontextualized conditions). Additionally, previous studies on vocabulary learning strategies have indicated that the most popular strategy among Japanese EFL learners is the rote memorization of word forms and translations (e.g., Maeda, Tagashira, & Miura, 2003; see also Mizumoto & Takeuchi, 2009). Thus, EFL learners' major source of vocabulary knowledge seems to be word forms and translations; the role of the reading context in intentional vocabulary learning has yet to be closely examined.

According to the previously mentioned studies, EFL learners appear to disregard contextual information given during intentional vocabulary learning. At least, it appears that the cognitive process of leaning new words would not be affected by the content of the context. In contrast to such an expectation, however, the five main experiments in Studies 1 and 2 indicated that Japanese university students engaged in vocabulary learning were actually sensitive to what was written in the context. In the series of experimental studies, participants were given a list of unfamiliar English words, their Japanese translations, and short verbal contexts and asked to learn them in a limited amount of time. To test whether the outcome of intentional vocabulary learning was influenced by contextual information, two types of contexts were used throughout the experiments: more and less imageable contexts. According to the results of Experiment 5, when two groups of high-intermediate learners with more and less imageable context were compared with respect to the scores on a gap-filling recall test, the more imageable group recalled the new words better than the less imageable group (see Section 4.5.3). Thus, because the learners' recall performance was enhanced by the imageable contexts, it was indicated that the learners utilized contextual information for

learning new words to some extent.

Given these results, one may argue that the two different types of contexts (i.e., more and less imageable contexts) caused different learning processes among the participants. This view can be further divided into strong and weak versions. The strong version assumes that the participants used the context carefully only when they felt the given contexts were imageable for them; otherwise, they disregarded what was written in the context. The weak version presumed that the participants read both types of context carefully, at least to some extent; however, they adopted different cognitive strategies while reading the contexts. These two expectations commonly assume that the learners' cognitive attention is mainly devoted to word forms and their translations even when contexts are provided. With respect to how learners would treat with more and less imageable contexts, both accounts seem to be plausible; the learners might apply either types of strategies in accordance with their teacher's instruction or learning environment. Since the present study explicitly asked participants to read contexts as well as word forms and translations (see Section 5.3.2.3), the weak version seemed to be more plausible than the strong version within the present study's paradigm.

The cognitive processes and strategies used in reading contexts have been a focus of theoretical approach in psycholinguistics and second language acquisition research. For example, Paivio's (1986, 1991, 2006) Dual Coding Theory assumes that the understanding of linguistic material involves two mental processes: language- and imagery-based processes. Applying this theoretical framework to the L2 reading context, Kadota (2006) called (a) the language-based process that handles the phonological, lexical, semantic, and syntactic features of language the *analytic process* and (b) the imagery-based process that enable learners to generate mental imagery and perform an inferential understanding of the context the *global process*. A possible expectation about the context imageability effect was that more imageable contexts would somehow induce learners' imagery-based (or global) process, whereas less imageable contexts would be comprehended mainly through language-based (or

analytic) process.

Considering the results from Studies 1 and 2, when highly imageable contexts are given to the learners, their mental process of context reading may include the generation of mental imagery. In contrast, when contexts that are not imageable are provided for the learners, their mental process of context reading might have a more analytic nature; for example, their typical cognitive strategy during context reading presumably includes syntactic parsing and access to lexical knowledge. However, to reveal whether the cognitive process of reading the context would actually be different between more and less imageable contexts, it was necessary to design a new experiment focusing on cognitive strategies employed during intentional vocabulary learning.

To examine what cognitive strategies EFL learners use while learning vocabulary, this study adopted the think-aloud approach. This technique is defined as a procedure requiring participants to verbalize and report the content of their thoughts while doing a task (Nassaji, 2003; see also Section 2.5.1). Despite some criticism that learners' think-aloud performance is not a true reflection of their cognitive process, think-aloud protocols often provide useful data about what is currently processed in learners' working memory. The most typical instruction in a think-aloud study is to ask participants to disclose their unedited thoughts while engaged in a given task (Yamashita, 2011). The present study gave this kind of instruction to participants who were learning a list of unfamiliar word forms, translations, and short contexts. The think-aloud methodology has been applied in both vocabulary learning (e.g., de Bot et al., 1997; Nassaji, 2006) and reading comprehension studies (e.g., Horiba, 2013; Kaakinen & Hyönä, 2005). However, within the research field of L2 vocabulary learning, this approach was mainly used to examine lexical inferencing strategies rather than intentional learning using translations (i.e., paired-associate learning). Therefore, there have been no established strategy taxonomies for a study of intentional vocabulary learning using word forms, translations, and contexts. Therefore, through Pilot Study 4, the present study carefully

created a new taxonomy by combining different taxonomies used in lexical inferencing and reading studies (see Section 5.2).

5.2 Pilot Study 4

5.2.1 Purpose

As introduced in Section 5.1.2, Experiment 6 was designed as a think-aloud study. However, the experimental procedures were different from the previous lexical inferencing studies because participants were asked to report their thought during reading not only English (i.e., L2) target words and contexts but also Japanese (i.e., L1) translations. Because of the presentation of L1 words, the experimental situation was quite different from past studies utilizing think-aloud methods. For example, as a lexical inferencing study, de Bot et al. (1997) asked 10 students to read a passage and collected think-aloud protocols about how they guessed the meaning of unfamiliar words. This study distinguished three major linguistic sources of lexical inferencing, as well as extralinguistic world knowledge and five minor sources. However, most of these knowledge sources might become unnecessary when the target word meanings are explicitly given to the participants. In contrast, Nassaji's (2006) taxonomy included not only identifying unfamiliar word meanings, but also evaluating and monitoring strategies; these strategies were expected to be observed in the present study's paradigm.

On the other hand, to examine how the analytic and global comprehension processes would be carried out during a reading context, it was necessary to further develop the think-aloud taxonomy in terms of context-based strategies. Kaakinen and Hyönä (2005), who examined L1 readers' comprehension processes, identified two deeper processes (i.e., self-explanations and questions), two shallower processes (i.e., associations and paraphrases), and monitoring. However, a pilot study indicated that participants frequently engaged in word- and sentence-level analysis while reading contexts (see Section 5.3.2.1), which were

not considered in Kaakinen and Hyönä's criteria. It seemed difficult to apply this taxonomy in an L2 study. In contrast, Horiba's (2013) taxonomy, which included structure analysis (i.e., word and sentence analysis), in-text inference (i.e., backward and predictive inference), reader response (i.e., association, reaction, and evaluation), and self-monitor, seemed to cover possible oral reports by the present study's participants. In this way, the final version of the strategy taxonomy for the present study was prepared based on different studies (see Section 5.3.2.4).

5.2.2 Method

5.2.2.1 Participants

The participants in the pilot study were three Japanese graduate students majoring in applied linguistics at a Japanese university, to whom the author fully explained the purpose of the study before the experimental treatments. Their data were not considered in the main analysis reported in this article; however, the author modified the experimental procedures and scoring criteria for the main study on the basis of these three students' responses and reports.

5.2.2.2 Materials

The learning materials used in this study were 10 target words paired with translations and verbal contexts, which were also used in Experiment 4 (see Appendix 3.6). The detailed description is presented in Section 5.3.2.2.

5.2.2.3 Procedure

The participants each took part in a series of tasks as follows: two practice sessions, contextualized learning, contextualized testing, and oral interview about experimental procedures. Before the two practice sessions, the purpose of the study was carefully explained

in Japanese. The overall procedures were similar to those in the main experiment (i.e., Experiment 6), which is explained in detail in Section 5.3.2.3. However, the mental imagery rating introduced in Section 5.3.2.3 was not conducted in Pilot Study 4. In addition, the method in Experiment 6 was improved based on the observation of the participants in Pilot Study 4. For example, the experimenter's instruction was clarified in wording in Experiment 6.

5.2.2.4 Scoring

As described in detail in Section 5.3.2.4, the participants' oral reports were recorded, transcribed and coded. Coding procedures included transcription of audio data, division into statements, and coding of learner strategies. Initially, the learner strategies were categorized based on Nassaji (2006) and Horiba (2013); however, according to the present study's purpose, the coding criteria were modified (see Section 5.3.2.4).

5.2.2.5 Analysis

The data were not analyzed because the Pilot Study 4 was conducted to establish a new taxonomy of the cognitive strategies that can be observed in the intentional learning of a word list with translations and contexts.

5.2.3 Results

Although the data were not analyzed statistically, Section 6.1.3 describes some typical protocols collected in Pilot Study 4 because it seemed necessary to consider what cognitive strategies more advanced learners would use to comprehend the contexts; the participants of Pilot Study 4 were graduate students majoring applied linguistics (see Section 5.2.2.1).

5.3 Experiment 6: Cognitive Strategies During Reading

5.3.1 Purpose

Experiment 6 aims to explore the mental processes involved in reading context during intentional learning of EFL vocabulary. Specifically, this study presents the results of an experiment designed to yield think-aloud protocols and two sets of supplementary data, namely, learning time and self-evaluations of imagery generation during learning. Think-aloud protocols were collected as each participant learned unfamiliar English words paired with Japanese translations and short verbal contexts. The data were analyzed in terms of (a) whether the learner's attention was mostly paid to word forms or to contexts and (b) by means of what kind of process she or he comprehended and utilized the contexts. Learning time data and imagery ratings were analyzed to supplement the findings from the think-aloud protocols. On the basis of these data, this study focused on how EFL learners' mental processes differ between situations where they can more easily generate mental imagery from contexts and those where this is less easy. In other words, this study compared more and less imageable contexts in terms of how the learners' cognitive processes differed.

The focus of this study was on how the cognitive process of intentional vocabulary learning differs according to context imageability. On the basis of previous studies by other researchers and of the author's own prior experiments, the following assumptions were made:

H6-1: In the intentional, translation-based mode of vocabulary learning, high-intermediate EFL learners' cognitive attention is mainly devoted to word forms and their translations, even when contexts are provided.

H6-2: When learners are provided with contexts that are not imageable, their mental process of context reading tends to be more analytic in nature.

H6-3: In contrast, when learners are given highly imageable contexts, their mental process of context reading tends to include the generation of mental imagery.

H6-1 was set on the basis of Webb (2007a), in which the university student participants seemed somewhat insensitive to contextual information, in that their major source of learning was the L1 translations. This result seems compatible with the findings of other studies (e.g., Maeda et al., 2003). On the other hand, H6-2 and H6-3 might appear somewhat inconsistent with H6-1, because these two hypotheses assumed some degree of learner sensitivity to context types. In fact, although H6-1 hypothesizes that the proportion of processes in a learner's mind that are context oriented is relatively small on the whole, H6-2 and H6-3 further imply that the learner's mental process of context reading should be characterized differently according to context imageability. As mentioned earlier, Kadota's (2006) model assumes that the cognitive process undergone by a learner engaging in context reading consists of two different but closely related sub-processes: (a) the analytic processes such as phonological encoding, lexical access, and syntactic parsing and (b) global processes including generation of mental imagery. H6-2 and H6-3 hold in common the assumption that the distribution of these two types of cognitive processes will be affected by the imageability of the context.

H6-1 and H6-2 (and, possibly, H6-3) were tested through analysis of learners' think-aloud protocols; for these, the participants were asked to report everything they thought while learning the unfamiliar word forms with translations and example sentences. Empirical support for H6-1 would be provided by the participants' dominant production of verbal protocols related to the word forms and translations rather than the contexts. H6-2 would be supported by the fact that less imageable contexts induce a higher frequency of analytic processes compared to the more imageable context condition.

In contrast to the above results showing support for H6-1 and H6-2, the results for H6-3 were more mixed. The participants might have difficulty verbalizing imagery-based reading processes while learning new words because imagery report itself requires cognitive efforts.

Therefore, after the learning phase, the author explicitly asked the participants if they imagined something in mind while they were reading the example sentences (i.e., the author elicited a *mental imagery rating*). The two possible cases taken to support H6-3 were (a) when higher imageability contexts induced more imagery-based think-aloud protocols and (b) when mental imagery ratings after learning were higher for the higher imageability contexts.

Through the investigation of these hypotheses, this study aims to explain how the learner's cognitive attention is distributed during intentional vocabulary learning. However, the frequency data do not clearly reflect the learner's degree of mental effort; in other words, one cannot tell whether the analytic processes are more demanding than the imagery processes, or otherwise. In addition to the hypotheses, an RQ was set as follows:

RQ6: Is the degree of mental effort different between more and less imageable contexts?

This study assumed that degree of mental effort during vocabulary would be reflected in and could be measured by the amount of learning time dedicated to target words with more imageable contexts as against those with less imageable contexts. Two different predictions were considered with regard to the RQ6: first, that learners would take longer for less imageable contexts—that is, that analytic processes would require greater mental effort by the learner than global processes; or, alternatively, that learning time would be longer for more imageable contexts—that is, that imagery generation during intentional vocabulary learning would be a more effortful process than the analytic processes. Grabe (2009) states that higher level reading processes such as effective use of background (world) knowledge and inference generation are more challenging for L2 learners than lower level processes such as word recognition and phonological, semantic, and syntactic processing (see Section 2.3.1). Given that the higher and lower level processes almost correspond to global and analytic processes, respectively, the second prediction arguing that the global (i.e., imagery-based, higher level)

processes require greater mental effort seemed as plausible as the first one.

5.3.2 Method

5.3.2.1 Participants

Nine Japanese undergraduates took part in the main experiment; none of them had participated in any of the author's prior studies. The participants were second- or third-year students majoring in linguistics or education who had been learning English as a foreign language for at least seven years. In Experiment 5, participants' lexical proficiency was assessed according to scores on the first subsection of the STEP Eiken test, which measures lexical skills on the basis of use of vocabulary in context. According to scores on a part of the same test, the current participants' lexical proficiency level was roughly estimated to be high intermediate in Experiment 5.

5.3.2.2 Materials

The learning materials used in this study were (a) 10 target words, (b) 5 more and 5 less imageable contexts, and (c) the lexical proficiency test (see Appendix 3.6). The target words and contexts were prepared based on Experiment 4, which also used STEP Eiken test materials (STEP, 2010).

Target words. Experiment 6 adopted real words that were assumed to be unfamiliar to the participants. There are both advantages and disadvantages to adopting this approach and rejecting nonwords (e.g., *denent*, used in Experiment 1 in the present thesis; adapted from Webb, 2007a). The most important advantage of this approach is that the experiment will resemble a real learning situation. In the intentional mode of vocabulary learning, participants might utilize the morphological features of a given word as learning support, or make analogies between the new word and already known words. If the experiment had used

nonwords such as *denent*, which was said to mean 'to remember' in Webb (2007a), this morphological analysis or analogy technique would fail to function to reinforce lexical knowledge. This was not considered desirable, since the goal was to compare the frequency of strategy use based on L2 word forms, L1 translations, and the analytic and global processes of reading context. On the other hand, there is also a defect in this approach, namely, that any prior knowledge of target words by learners cannot be controlled. It is possible that some participants might have had partial knowledge of some target words at the time of learning. However, it should be emphasized that even if the list did include partially known words as well as perfectly unknown words, the meanings of all the target words were provided for the participants and all the contexts were new to them.

Contexts. Contexts were prepared for the 10 target words based on STEP Eiken test materials (STEP, 2010), as also employed in Experiment 4; they were modified slightly for the present study. The types of discourse in which they were presented (that is, monologue or dialogue) were not changed from the original. The number of words per context varied from 17 to 40 (M = 26.70, SD = 7.87). According to calculation by Microsoft Word 2010, the readability measures were as follows: Passive Sentences = 0.0%, Flesch Reading Ease = 62.5, Flesch-Kincaid Grade Level = 8.2. The contexts were divided into more and less imageable groups. The division was made based on imageability ratings on a seven-point Likert-type scale performed in a norming study (N = 10). As described in Section 4.2.3, this norming study further confirmed a high reliability of the imageability rating by *test-retest method*, where the same participants performed the same ratings (but in a shuffled order of item presentation) three weeks after the first session (r = .91).

5.3.2.3 Procedure

The nine participating students each took part in a series of paper-based tasks for each

of several procedures: the proficiency test, two practice sessions, contextualized learning, and contextualized testing. Before the two practice sessions, the purpose of the study was carefully explained in Japanese, as follows: “The experiment aims to examine how students learn new English words using a word list with example sentences. You will learn 10 words in around 10 to 15 minutes. Each item has a Japanese translation and a context. I would like you to report what the process of your learning was like.” In case the participants got confused about what they were going to do, the procedure of the think-aloud task was not explained in detail until the first practice session was completed.

Practice sessions. Because it seemed challenging for EFL learners to verbalize their thoughts while learning unfamiliar words, the participants needed to be familiarized with the think-aloud procedures in a step-by-step manner before the main learning phase. The author first presented them with a sample English word (not included in the main word list) together with its translation and an example sentence. The participants were given 30 seconds to read this material silently and learn the spelling, meaning, and use of the target word. Then, the author initiated the interview, as follows: “Now, I will ask you some questions about how you learned this. First, when you started learning, which did you look at first: the word form, translation, or context?” If a participant answered that she or he saw the word form(/translation/context) first, the author asked what she or he thought, felt, or imagined when she or he looked at it.

After fully explaining their retroactive mental processes of learning in the first practice session in this way, the participants were informed that they would need to verbalize everything they thought or felt while learning some more new words. To familiarize them with the think-aloud procedure, two new English words paired with their translations and contexts were presented to the participants, and they were asked to report everything about the material they were looking at and what they were thinking as they learned the target

word's spelling, meaning, and use. This series of practice procedures was developed on the basis of the results of the pilot study with the three graduate students (see Section 5.3.2.1).

Contextualized learning (think-aloud). In the learning phase, the participants were asked to learn a list of 10 target words in context along with their meanings in Japanese. Because the tester's intention was fully explained and the participants' attention was directed to memorizing the list, this task can be regarded as an intentional learning task (Hulstijn, 2005; Schmitt, 2010). During the learning phase, the participants were also asked to report everything they thought and felt, just as in the practice session. All participants had agreed to their verbal reports throughout the learning phase being recorded by digital recorder. Although the author had previously announced that the learning phase was expected to last for 10 to 15 minutes, no hard time limit was specified. Actual learning time was approximately 10 minutes on average (M = 604 s, SD = 191 s, Min = 338 s, Max = 922 s); detailed analysis of the results by learning time is done in the "Time Allocation" section below.

After learning, the list was removed and a contextualized recall test using a gap-filling approach was conducted, in which learners needed to remember the word form matching a given context. The purpose of conducting the recall test was to create a situation where the participants made an effort to learn all 10 target words. The contexts presented were the same as those in the learning list. The recall test had no time limit and was finished in approximately six minutes on average (M = 358 s, SD = 208 s, Min = 173 s, Max = 879 s).

Mental imagery rating. Finally, the participants again received the list of 10 words, translations, and contexts. They were asked to think back to how they had learned them and report whether or not they had called up any mental imagery from each context during learning. A "Yes" answer was coded as 1, "No" as 0. When the participant thought she or he might have imagined some partial representation but was not quite sure about it, the answer

was labeled as 0.5. The rating format is presented in Appendix 5.1.

5.3.2.4 Scoring

The participants' oral reports were transcribed and coded, with reference to Yamashita (2011). Coding procedures included (a) transcription of audio data, (b) division into statements, and (c) coding of learner strategies; Table 5.1 shows examples of them.

Table 5.1

Examples of Transcription and Coding of the Verbal Reports

Step	Example
Audio data	/əp/ /əprentɪs/ (short pause) /dɪskre/ (longer pause) /ə:/ /dɪskredət/
Transcription	App-、apprentice。Discre-。(2 s) あー discredit。
Division	(1) App-、apprentice。/(2) Discre-。(2 s)/(3) あー discredit。
Coding	(1) 01A, (2) 02B, (3) 02B

Note. For transcription, pause length was recorded for pauses of two seconds or more; Japanese punctuation marks were added for the coder's easy reading; Japanese word "あー" is just a filler like *uh* or *aha*. In the division step, the above example identifies three statements, (1), (2), and (3). In the coding step, each statement was replaced with the labels representing the appropriate strategy categories in Table 5.2; the numbers here represent target items.

Transcription of audio data. First, the audio data were transcribed in a combination of Japanese and English, as seen in Table 5.1; the participants mainly used Japanese during the think-aloud phase but often read the target words and contexts aloud in English. Punctuation marks were used in these transcriptions for easier reading. Pause length was noted for pauses of two seconds or more.

Division into statements. Next, the transcript was parsed into *statements*, which were roughly equivalent to clauses (Horiba, 2013), and places each statement into one of several reading strategy categories. As shown in Table 5.1, the present analysis used slash marks to indicate statement divisions at (a) the boundary between two clauses, and (b) the boundaries of responses to different target items, and (c) a pause of two seconds or more. All these criteria for statement division were determined on the basis of the results of the pilot study.

Coding of learner strategies. After division into statements, each statement was coded for strategy categories. Table 5.2 summarizes the 11 strategy categories applied; these categories were set mainly based on Nassaji (2006) and Horiba (2013). The 11 categories were sorted into four groups as follows. First, the *Repetition* group involved reading aloud an English word and a Japanese translation (i.e., *word form* and *translation* strategies) as well as a context. For reading-aloud of context, reading of sentences including a target word was distinguished from reading of other sentences (i.e., *local* and *global context* strategies, respectively). Second, the *Analysis and Analogy* group included use of knowledge about English word parts and part of speech (i.e., *L2 word analysis*), and synonyms of a target word or its translation (i.e., *L2* and *L1 word analogy* strategies, respectively). In addition, this group included syntactic analysis and verbatim translation of a sentence (i.e., *sentence analysis*). The strategies concerning context comprehension, that is, repetition of local/global context and sentence analysis, were considered analytic processes in terms of the reading comprehension model. To consider reading at a deeper level and capture global (or imagery-based) processes, this study also set the third strategy group *Inferential Comprehension of Context*, which included any reports that referred to something not explicitly written in the context (i.e., *implicit idea*) and mental images (i.e., *imagery report*). Learners sometimes also mentioned whether they felt that a target item was easy or difficult to learn; these reports were labeled *Self-Monitoring* (or *monitoring* strategy). Finally, there were

some cases where learners' verbal protocols did not fit in any of the 11 strategy categories, for example, when a participant counted how many items the word list had. These kinds of statements were labeled *other*.

Table 5.2

Simplified Definitions and Examples of the Cognitive Strategies

	Category	Definition	Example of think-aloud protocols	Knowledge source
		Repetition		
A	Word form	Reading a target word form	"*Discredit*, *discredit*."	Word-based (L2)
B	Translation	Reading a translation	"*Hyoban o waruku-suru*."	Word-based (L1)
C	Local context	Reading a sentence including a target word	"*An attempt by his critics to discredit him.*"	Context-based (analytic)
D	Global context	Reading a sentence separated from a target word	"*Rumors that the presidential candidate...*"	Context-based (analytic)
		Analysis and Analogy		
E	L2 word analysis	Using a morphological clue	"The prefix *dis*- means something negative."	Word-based (L2)
F	L2 word analogy	Referring to a word related to the target word	"Is this word related to credit card?"	Word-based (L2)
G	L1 word analogy	Referring to a word related to the translation	"*Akuhyo o tateru.*"	Word-based (L1)

Table 5.2 (*continued*)

	Category	Definition	Example of think-aloud protocols	Knowledge source
H	Sentence analysis	Parsing and translating a sentence	[Verbatim translation of a sentence]	Context-based (analytic)
		Inferential Comprehension of Context		
I	Implicit idea	Understanding an implicit idea	"Oh, the candidate had a rival."	Context-based (inferential)
J	Imagery report	Imaging something from an example sentence	"I can imagine his feeling."	Context-based (inferential)
		Self-Monitoring		
K	Monitoring	Reporting some ease or difficulty learning	"This word is very difficult to learn."	N.A.

Note. Statements that did not fit in any of the 11 categories above were labeled as *other*. Although the 11 categories were initially grouped into *Repetition*, *Analysis and Analogy*, *Inferential Comprehension of Context*, and *Self-Monitoring*, these groups do not represent where the learner is paying attention. Therefore, with regard to the knowledge source, this study discriminated between the word- and context-based strategies.

It might be worth noting that each of the 11 categories was further specified for individual target items. Taking the example of repetition of word forms, the number of repetitions of *apprentice* was counted independently from repetitions of *discredit* and other target items. Thus, 111 distinct learner strategies (10 target words × 11 strategy categories + the *other* category) were identified. One might suspect that it would be puzzling to classify each statement into only one of 111 subtypes; however, when two raters independently coded transcriptions from two participants in the main experiment, their coding showed 97.46%

agreement (that is, 9 disagreements out of 355 statements) when a few careless mistakes were disregarded. The two raters solved all their disagreements through discussion

5.3.2.5 Analysis

Strategy frequency. To examine whether participants paid cognitive attention to word forms and translations, this study compared the total number of word- and context-based strategies evidenced using a chi-square test. Namely, the aggregate subtotal for word form (A), translation (B), word analysis (E), L2 word analogy (F), and L1 word analogy (G) was compared to the subtotal of local context (C), global context (D), sentence analysis (H), implicit idea (I), and imagery report (J) strategies. The self-monitoring strategies (K) were not considered in this calculation because it was difficult to specify word-based and context-based monitoring. In addition, another chi-square test was conducted to examine whether the frequency of word- and context-based strategies had a relation with context imageability.

However, the above analysis did not distinguish the types of context-based strategies. Therefore, to assess the effect of context imageability on learners' cognitive attention during vocabulary learning, context-based strategies were divided into those involving analytic and inferential processes. A chi-square test was carried out for a cross-tabulation between context type (i.e., more/less imageable) and the context-based process (i.e., analytic/inferential). For all chi-square tests, the alpha level was set at .05.

Mental imagery ratings. The mental imagery ratings were based on each participant's report regarding whether or not she or he called up mental imagery from a given context during learning (1 = called up imagery; 0.5 = called up partial imagery; 0 = did not). The average ratings were compared between the more and less imageable conditions using a *t* test; the alpha level was set at .05.

Time allocation. The audio data were analyzed in terms of how long the participants learned each item. The total amount of time spent for the five more imageable items was compared with the five less imageable items using a *t* test; the alpha level was set at .05.

5.3.3 Results

Three types of data were obtained in this study: (a) think-aloud protocols, (b) mental imagery ratings, and (c) time allocation data. First, think-aloud protocols were analyzed to examine the relative frequency of use of different strategies (addressing H6-1 and H6-2). Second, the mental imagery ratings were analyzed to test imagery generation from context (H6-3). Third, time allocation was analyzed to examine the allotment of mental effort (RQ6).

Strategy frequency. As just mentioned, relative frequency of participants' strategy use was analyzed in order to examine H6-1 and H6-2, which argued that EFL learners' cognitive attention is mainly devoted to word-based processes (that is, word forms and translations), while the nature of context-based processes (analytic and inferential processes) is more influenced by context imageability. Tables 5.3 and 5.4 abstract the raw frequency data in order to test H6-1 and H6-2 (see also Appendix 5.2).

First, as explained in Section 5.3.2.5, the total number of word- and context-based strategies was compared (i.e., the aggregate subtotal for A, B, E, F, and G vs. the subtotal of C, D, H, I, and J). The result showed that participants used word-based strategies more frequently (627 times in all) than context-based strategies (397 times; see Table 5.4), $\chi^2(1) = 51.66$, $p < .001$. In fact, most participants first repeated all 10 word forms associated with translations but sometimes skipped reading-aloud or making comments on contexts. Thus, H6-1 was fully supported in this study.

Table 5.3

Frequency of the Cognitive Strategies as a Function of Context Imageability

	Category	Knowledge source	More imageable condition	Less imageable condition	Total
A	Word form	Word-based (L2)	167 (14.60%)	162 (14.16%)	329 (28.76%)
B	Translation	Word-based (L1)	132 (11.54%)	117 (10.23%)	249 (21.77%)
C	Local context	Context-based (analytic)	113 (9.88%)	155 (13.55%)	268 (23.43%)
D	Global context	Context-based (analytic)	29 (2.53%)	35 (3.06%)	64 (5.59%)
E	L2 word analysis	Word-based (L2)	11 (0.96%)	23 (2.01%)	34 (2.97%)
F	L2 word analogy	Word-based (L2)	6 (0.52%)	8 (0.70%)	14 (1.22%)
G	L1 word analogy	Word-based (L1)	1 (0.09%)	0 (0.00%)	1 (0.09%)
H	Sentence analysis	Context-based (analytic)	15 (1.31%)	32 (2.80%)	47 (4.11%)
I	Implicit idea	Context-based (inferential)	9 (0.79%)	3 (0.26%)	12 (1.05%)
J	Imagery report	Context-based (inferential)	5 (0.44%)	1 (0.09%)	6 (0.52%)
K	Monitoring	N.A.	26 (2.27%)	31 (2.71%)	57 (4.98%)
-	Other	N.A.	N.A.	N.A.	63 (5.51%)

Table 5.4

Frequency of Word- and Context-Based Strategies as a Function of Context Imageability

	Word-based			Context-based			
Image-ability	L2 (AEF)	L1 (BG)	Subtotal (AEF + BG)	Analytic (CDH)	Inferential (IJ)	Subtotal (CDH + IJ)	Total
More	184	133	317	157	14	171	488
Less	193	117	310	222	4	226	536
Total	377	250	627	379	18	397	1024

Note. The letters from A to J correspond to the strategy labels given in Table 5.2.

When the context imageability factor was taken into account, it was found that the frequency of word- and of context-based strategies had a significant relation with imageability, $\chi^2(1) = 5.46$, $p = .019$. As shown in Table 5.4, more context-based strategies were observed for the less imageable context. However, this result becomes much easier to interpret when the results of the next analysis are considered.

Second, to assess the effect of context imageability on learners' cognitive attention during vocabulary learning, context-based strategies were divided into those involving analytic and inferential processes. The result of a cross-tabulation between context type (i.e., more/less imageable) and the context-based process (i.e., analytic/inferential) showed a significant relation between these two variables, $\chi^2(1) = 9.26$, $p = .002$. Table 5.4 indicates that the frequency of the analytic processes was higher for the less imageable contexts than for the more imageable ones, which supported H6-2. The relationship between context imageability (more/less) and word-based process (L2/L1) was not significant, $\chi^2(1) = 1.16$, $p = .281$.

On the other hand, more inferential processes were found for the more imageable contexts than for the less imageable ones. This result seems initially to support H6-3, the

hypothesis that imageable contexts would enhance learners' imagery-based comprehension. However, all 18 statements labeled as inferential processes of context were produced by only two participants (S7 and S8). Therefore, it was safer to discuss H6-3 based on mental imagery rating in the next section.

Mental imagery ratings. As a result, the probability of mental imagery generation was significantly higher for the more imageable contexts (M = 0.70, SD = 0.22) than for less imageable contexts (M = 0.56, SD = 0.19), $t(8) = 2.60$, $p = .032$ (see Appendix 5.3). Together with the strategy frequency data reported earlier, these results suggest that the mental process undergone by EFL learners reading imageable contexts includes generation of mental imagery, supporting H6-3.

Time allocation. The five items using low-imageability contexts (M = 297.11 s, SD = 171.29 s) required approximately ninety seconds more processing time than the high-imageability ones (M = 206.44 s, SD = 105.77 s; see Appendix 5.3); this difference was statistically significant, $t(8) = 2.91$, $p = .020$. It is worth noting that the total amount of time is not a simple sum of the time spent on the more and on the less imageable items; instead, the total time includes thinking time that could not be reasonably interpreted as relating to learning of a specific item (e.g., a participant's counting the items on the word list).

Based on these results, it seems that the EFL learners allotted different amounts of mental effort between the more and the less imageable contexts. Considering this finding in relation to the think-aloud and mental imagery rating data, it seems highly possible that the less imageable contexts required the use of analytic processes by the learners, and that this approach was effortful enough to diminish their cognitive resources available for the global, imagery-based processes. However, these results should be interpreted only cautiously, because the time allocation data were not always directly reflective of learners' cognitive

effort. Within the framework of this study, learning time on each item could be greatly affected by how many oral reports the participants made. A more reasonable suggestion might be that the more imageable contexts, which induced generation of mental imagery, were not as demanding to comprehend as the less imageable ones were.

5.3.4 Discussion

This experiment achieved three major findings on the basis of the think-aloud protocols, learning time data, and self-evaluations of imagery generation during learning. The current section summarizes the findings and interprets them in light of the characteristics of the mental processes of EFL learners involved in intentional vocabulary learning. In Chapter 6, the results are reviewed for pedagogical implications and integrated into a working model of cognitive processes during contextualized vocabulary learning.

With regard to the first finding, this study hypothesized that participants' attention would mainly be devoted to the word forms and the translations even when contexts were provided (H6-1). This hypothesis was supported by the result showing that the frequency of participants' use of word-based learning strategies, which corresponds to their attention to L2 word forms and L1 translations, was considerably higher than that of context-based strategies. This suggests that the most typical strategy among these university students with high-intermediate L2 English lexical proficiency was to repeatedly associate the L2-L1 translation pair when learning new words. This finding is compatible with past studies showing that repetition strategies are the dominant strategies used by Japanese students to learn English vocabulary (e.g., Maeda et al., 2003), although as far as the present author is aware, no past studies have reported what Japanese EFL learners were actually doing during intentional vocabulary learning using translations and contexts. However, with regard to the current findings, the importance of the frequency of context-based learning strategies should not be underestimated. In fact, it was the context-based strategies that made a more notable

difference for individual learners. Therefore, teachers should monitor and guide how their students use contexts while learning vocabulary.

The second finding was that the learners used more imageable contexts differently from less imageable contexts. The results supported the proposition that the participants' mental process of context reading would be more analytic in nature when less imageable contexts were provided (H6-2). On the other hand, when they were given more imageable contexts, their process tended to include generation of mental imagery (H6-3). Taken together, these findings imply that teachers should carefully consider what type of context to present to their students. If teachers want their students to practice analytic sentence comprehension, it could be helpful for them to present the students with less imageable contexts, which can hardly be understood solely by top-down, inferential comprehension processes. However, according to the results from Studies 1 and 2, teachers should use imageable contexts if they want to create opportunities for vocabulary learning that lead the learners to focus on the content of contexts.

Third, as to the degree of mental effort devoted to the contexts (RQ6), the results indicated that less imageable contexts required greater effort from learners. From a pedagogical point of view, Experiment 5 found that imageable contexts enhanced vocabulary learning among high-intermediate EFL learners. These findings, in combination with the results in Experiment 6, suggest that students spend more time learning less imageable items but learn more imageable items better. Experiment 6 supplies evidence that the more imageable contexts enable learners to call up a mental image for a vocabulary item being learned. Additionally, it was found that this process of imagery generation was not cognitively demanding for high-intermediate learners. On the other hand, during learning of less imageable contexts, the participants performed analytic comprehension processes more frequently than during learning of imageable contexts; this might have been a burden for the L2 learners, who could not read the written materials as fluently as L1 readers. This seems compatible with Laufer and Shmueli's (1997) suggestion that it might be inefficient for L2

learners to read a long passage and that the minimal length of context was the best.

Given the findings summarized above, we see that imageable contexts require less effort and are easier to utilize than less imageable contexts. However, another expectation was held before conducting this study, namely, that the students would exert greater mental effort on the more imageable contexts in order to achieve more sophisticated comprehension. In the context of this expectation, it is necessary to discuss why imageable contexts induce better learning in a shorter time. As covered in Section 5.1.2, researchers often explain the mental process of L2 reading in terms of two different mental processes: Paivio (1986, 1991, 2006) hypothesized the interaction of the imagery-based and language-based processes; Grabe (2009) distinguished the higher and lower level processes; Kadota (2006) assumed different roles of the global and analytic processes. Although the descriptions of these processes were quite different, one common assumption among these views might be that the inferential processes activating the reader's prior knowledge plays a different role from the verbatim and word-by-word decoding of sentences. However, one of the most basic assumptions of Dual Coding Theory is that the most effective learning and understanding occur when the language-based (i.e., analytic) component is activated simultaneously with the imagery-based (i.e., global) component. Another important assumption of this model is that the global processor works much more quickly, or in other words, automatically compared to the time-consuming processing of the analytic processor. These assumptions are quite consistent with the current finding that the more imageable contexts induced better learning in a shorter time than the less imageable contexts.

Based on the present study, the model can be further developed in terms of how the language-based (i.e., analytic) and the imagery-based (i.e., global) processors work during intentional learning of L2 vocabulary using translation and contexts. First, it can be assumed that learners' cognitive attention during intentional vocabulary learning is mainly devoted to word forms and translations; therefore, learners comprehend contexts with only a limited

amount of working memory capacity. To optimize working memory performance, the learner's cognitive attention is oriented toward either analytic or global processes according to the imageability of a given context. In Studies 1 and 2, university students at a high-intermediate English proficiency level more target words when imageable contexts were given as cues; in addition, according to Pilot Study 1, when given target words as cues, they recalled the content of more imageable contexts better than that of less imageable contexts. The current model's assumptions are quite consistent with these results. However, researchers suggest that the intentional vocabulary learning using translations accounts for only a limited part of the L2 lexical development (Nation, 2013). Therefore, further research is needed to explore how the above assumptions can be applied in various mode of contextualized vocabulary learning such as meaning-focused input activities.

5.4 Summary of Study 3

To explore the mental processes involved in reading example sentences to learn EFL vocabulary, Experiment 6 was designed to yield think-aloud protocols and two sets of adjunct data, namely (a) learning time and (b) self-evaluations of imagery generation during learning. The following Hs and RQ were addressed:

H6-1: In the intentional, translation-based mode of vocabulary learning, high-intermediate EFL learners' cognitive attention is mainly devoted to word forms and their translations, even when contexts are provided.

H6-2: When learners are provided with contexts that are not imageable, their mental process of context reading tends to be more analytic in nature.

H6-3: In contrast, when learners are given highly imageable contexts, their mental process of context reading tends to include the generation of mental imagery.

RQ6: Is the degree of mental effort different between more and less imageable contexts?

Among the above statements, three Hs were tested through the analyses of think-aloud protocols, and the RQ was examined based on learning time. The self-evaluations of imagery generation were used to support the discussion of H6-3. During intentional vocabulary learning using translations, example sentences were processed with a limited amount of cognitive resources (Laufer & Shmueli, 1997). However, as suggested in Studies 1 and 2, the mental process of comprehending example sentences differed across learners and materials. In particular, when learners were provided with example sentences that were not imageable, their mental process of reading was more analytic in nature. In contrast, when highly imageable sentences were given to learners, their reading process tended to include the generation of mental imagery.

Test results revealed that each of the three Hs were supported. With respect to RQ6, this study obtained partial evidence suggesting that the learning processes related to context-based mental imagery were not as cognitively demanding as the analytic processes of comprehending context. In addition, these results may explain why imageable contexts were found to enhance learner recall of new words in Studies 1 and 2. Namely, the imageable sentences enabled learners to comprehend contextual information effectively using global and inferential processes. According to Paivio's Dual Coding Theory (1986, 1991, 2006), reading comprehension involves two mental processes: a global, imagery-based process and an analytic, language-based process. An important assumption of this model is that global processing occurs much more quickly than analytic processing. This assumption is consistent with the current finding, wherein the more imageable example sentences induced improved learning in a shorter timeframe.

To the author's knowledge, no past studies have closely examined how EFL learners read example sentences during the deliberate memorization of a word list, while focusing on the types of example sentence provided. Presumably, because recent studies have suggested that learners are unable to utilize contextual information effectively, the role of context in

intentional vocabulary learning has been underestimated. However, as far as high-intermediate learners are concerned, it seems important for them to understand the meanings of L2 words in context through activation of their background knowledge and the generation of mental imagery, even when the word meanings themselves are presented as L1 translations. However, one of the current experiment's limitations was that the focus was too narrow, and its sample size was too small to generalize the findings. Future studies should further examine whether this study's suggestions can be applied to other learning situations, including incidental vocabulary learning through extensive reading.

Chapter 6

General Discussion

6.1 Three Major Findings

To reveal the function of context reading in intentional vocabulary learning and explore the effective use of example sentences, this current investigation conducted six experimental studies. This chapter first summarizes the major findings of the experiments and then discusses more theoretical matters. As illustrated in Table 6.1, the research findings are overviewed and discussed from the following three viewpoints: (a) sensitivity to contextual information (Section 6.1.1), (b) interaction between imageability and proficiency (Section 6.1.2), and (c) different processes of context reading (Section 6.1.3).

Table 6.1

Summary of the Six Experiments, Three Viewpoints, and Two Focuses in This Study

Experiment	Viewpoint	Focus 1: Condition	Focus 2: Process
Experiments 1, 2, and 3 (Study 1)	(a) sensitivity to contextual information	More imageable contexts may enhance vocabulary learning.	Imageable contexts are better remembered in relation to new words.
Experiments 4 and 5 (Study 2)	(b) interaction between imageability and proficiency	The context imageability effect is selective about learner proficiency.	Poorer learners cannot take advantage of mental representations.
Experiment 6 (Study 3)	(c) different processes of context reading	Effective use of contexts is related to inferential understanding.	Learners read more and less imageable contexts differently.

As shown in Table 6.1, each viewpoint corresponds to each study group in principle, as follows: The first viewpoint corresponds to Study 1 (Experiments 1, 2, and 3); the second to Study 2 (Experiments 4 and 5); and the third to Study 3 (Experiment 6). For each viewpoint, the two main focuses of the current research are discussed. These focuses, introduced in Section 1.2, are as follows: (a) how teachers and learners can optimize the effect of context reading for vocabulary learning and (b) the nature of learners' comprehension process of context while their intention was oriented toward memorizing lexical items. The general discussion also considers the mental connection between lexical items and context information; as introduced in Section 1.2, after learners encounter lexical items in context, they might retain the contextual information in relation to the target lexical items.

6.1.1 Sensitivity to Contextual Information

Past studies focusing on contextualized vocabulary learning suggested that rote memorization of a word list is the most efficient way of learning the meanings of new words, when the learning rate within a limited time is the measure of learning (see Folse, 2004). Researchers agree that L2 learners tend to depend on translation-based recognition when they first encounter new words; this concept is also consistent with assumptions proposed by theories such as Revised Hierarchical Model. In fact, when Webb (2007a) compared vocabulary learning using a list of word forms and translations (i.e., decontextualized) and a list of word forms, translations, and example sentences (i.e., contextualized), he found no overall difference between the two conditions. Although his data apparently showed that context may have a small effect on receptive knowledge of syntagmatic association, he concluded based on careful analyses that a single context may only have a slight effect on collocational knowledge (i.e., syntagmatic association, in his terminology), which may be outweighed by other factors. This might be regarded as evidence showing that L2 learners are not quite sensitive to contextual information; it even seems they tend to ignore contexts. At

least, it was found that context presentation was not as effective as people have believed.

However, the present study found evidence showing L2 learners are more sensitive to context type than expected. If their learning took place in the same manner regardless of the presented contexts, their vocabulary test scores after learning would not be different across context types. In fact, their scores were different between the more and less imageable contexts, showing their sensitivity to context type to some extent; this phenomenon is closely connected with learner proficiency, but the interaction between context type and learner proficiency is discussed in the next section (Section 6.1.2).

It might seem surprising that this pattern of results was found repeatedly with different types of materials and experimental designs. For example, the context imageability effect was found with different target words including pseudowords (e.g., *denent*, which means "to remember"), rare words (e.g., *reminisce*), and words that were not too difficult or easy for learners (e.g., *clutch*). Also, this study used various types of contexts, from original material based on past studies and example sentences from a dictionary with some modification, to contexts that were originally developed for a vocabulary test. Regarding the experimental design, Experiments 1, 2, 4, and 6 treated the context imageability as a within-participant factor; namely, each participant received a list of words to remember, including both more and less imageable contexts. Experiment 3 also used a within-participant design, but the procedure was different: The participants learned different lists in different weeks to experience all three experimental conditions (i.e., more and less imageable context and no context). On the other hand, Experiment 5 compared a group of learners with a more imageable list and another group with a less imageable list, making the context imageability a between-participant factor. Based on a series of experiments all suggesting the effect of context imageability in some way, there seems to be enough evidence to argue that context reading can influence L2 learners' intentional vocabulary learning, at least university students at a proper proficiency level (see Section 6.1.2).

However, one might argue that the imageability effect found in the present research was caused by another context-related factor: semantic informativeness. As mentioned in Section 2.4.1 (see Table 2.2), context informativeness indicates how useful context information is in guessing an unfamiliar word (e.g., Chaffin et al., 2001; Godfroid et al., 2013; Webb, 2008); this factor is also called context pregnancy (e.g., Lawson & Hogben, 1996; Read, 2000) and contextual richness (e.g., Joe, 2010; Mondria & Wit-de Boer, 1991; Zahar et al., 2001). Although an imageable (or concrete) sentence is not always informative in terms of the guessability of a certain word used in that sentence, it is important to confirm that the current results were caused by context imageability, not informativeness. For example, the sentence *She asked a lot of people a lot of questions and always* <u>*faddamed*</u> ("wrote") *the answers in her notebook* is not only imageable, but also informative (Webb, 2008; see Appendix 3.1); that is, it is easy to guess the meaning of the target word. If the more and less imageable contexts used in the resent study were biased in terms of their informativeness, it would be quite difficult to determine whether the learners were sensitive to context imageability or informativeness.

To confirm that the contexts used in the present study were not biased in terms of their informativeness, an additional analysis was conducted using the Latent Semantic Analysis (LSA) tool (http://lsa.colorado.edu/). This is a method of extracting and representing context-based meaning and word usage by means of statistical computations applied to a large corpus of text (Launder, Foltz, & Laham, 1998). Using this tool, semantic overlap between the target word and contexts can be calculated; that is to say, LSA suggests how the target word in each context is semantically related to the complete contextual meaning. Calculation was performed using the "general reading (up to first year college)" and the "term to term" comparison method, in conjunction with the context materials presented in Appendices 3.4, 3.6, 3.8, and 4.4. The pseudowords (e.g., *faddam*, *hodet*, etc.) were replaced with synonyms (e.g., *write*, *face*, etc.), however among the 30 items *indite*, *reminisce*, and *esplanade* were not

analyzed, since they were not registered in the LSA corpus.

Semantic overlap between the target words and their contexts, which represented context informativeness, was not different between the more imageable (n = 27, M = .29, SD = .19) and less imageable contexts (n = 27, M = .30, SD = .19; p = .955). The complete LSA results are provided in Appendix 6.1 along with keyword abstraction results, which were calculated using AntConc (see Section 2.4.1). Similarly, Pilot Study 1 showed that the informativeness and imageability ratings had no significant correlation (ρ = –.03, p = .867). Therefore, as far as the present materials are concerned, the context effects found in this study were caused by the imageability factor. In fact, the guessability variables (i.e., context informativeness, pregnancy, richness, and constraint) might not be substantially influential in relation to the learning mode that the present study focuses on. Because the study's participants were engaged in an intentional and translation-based mode of learning (see Section 2.2.1), it was unnecessary for them to guess the target word meanings. Instead, the contextual information might have functioned to enrich the conceptual representation of the target words. When contexts are presented in combination with translations of target words that explicitly convey word meaning, the imageability of the contexts might be more important for learners than the informativeness.

In relation to the mechanism of how context imageability affects L2 vocabulary learning, Paivio's Dual Coding Theory predicts imageable <u>contexts</u> can be understood and learned better than less imageable ones. However, it was previously unclear whether the imageability of context affects the learnability of <u>new words</u>. That is, it was necessary to discuss a kind of transferability of the imageability effect from contexts to words. With respect to this problem, the results of Pilot Study 1 seem highly suggestive. The results of this experiment showed that when learners read a list of imageable contexts including new words and they were presented with the new words, they could recall in what context each target word was embedded. In other words, imageable contexts are better remembered in relation to

new words. This kind of memory of word-context association might be related to the concept called “ensemble” in the ICE model (Murnane et al., 1999; see also Franco-Watkins & Dougherty, 2006). However, the experiments in the present study were not designed to examine how meaningful word-context associations are. It should be interesting to investigate the meaningfulness of the relationship in future research; one may expect that more imageable contexts induce more a meaningful relationship between the new word and the context because Experiment 6 indicated a few learners made inferences about the reference of the target word when they were reading an imageable context. As assumed in Kintsch’s framework, readers retain what they read longer if they have constructed a mental representation at a higher level (van Dijk & Kintsch, 1983; see also Grabe, 2009; Tapiero, 2007). If the imageable contexts used in the present study led to richer representations in the learners’ minds, those mental representations might be connected to the information of the new word.

However, it is worth noting that this kind of context-dependent memory does not guarantee acquisition of full lexical knowledge. Researchers agree that learning, including vocabulary learning, follows some steps from “context-dependent” to “context-independent” memories. The term “context” here refers to a broader sense than example sentences for vocabulary learning; it covers any kind of environmental information such as what kind of paper was provided in an experiment. To avoid confusion, this term always accompanies quotation marks in this thesis when it is used in the broader sense. As overviewed in Kadota, Noro, Shiki, and Hase (2014), input information is firstly encoded with the “context” in mind and becomes what is called episodic memory. This memory is gradually “decontextualized” through learners’ experiences of encountering it in various “contexts.” With respect to L2 vocabulary learning in an English classroom, a new word is encoded in a learner’s memory with certain kinds of “contextual” information such as who taught him or her the word, in which lesson in the textbook he encountered the word, and in what kind of sentence the word

was used (i.e., episodic memory). However, after several months, she or he might forget these kinds of "contextual" information, and her or his knowledge of the word is thus "decontextualized." Finally, when the word is mastered, the learner will be able to use it without remembering any "contextual" information such as the teacher, textbook, or sentence (i.e., procedural memory; pp. 45–53).

Nevertheless, this issue does not mean that L2 vocabulary should be taught apart from context in a narrower sense (i.e., example sentences). Rather, the assumption of memory development from episodic to procedural knowledge might further emphasize the importance of presentation of context; if students learn a list of new words without any contextual information, their episodic memory of their first encounter with these words would lack information about linguistic context. As researchers have pointed out, vocabulary learning is an incremental process (e.g., Schmitt, 2010). The learner's lexical knowledge should be "decontextualized" through a number of encounters with words in various contexts; this kind of elaboration of knowledge through repeated encounter is also referred to as *relearning* (Fukkink, Blok, & de Glopper, 2001). An important suggestion from the current research is that L2 learners are actually able to encode contextual information while they are engaged in learning new word forms and meanings. Their memory of context is associated with the lexical representation in their minds even if they have learned the target words only once in a short time. However, construction of the word-context association is closely related to learner proficiency. Section 6.1.2 explains the relationship between the word-context association and learner proficiency as another important finding of this study.

6.1.2 Interaction Between Imageability and Proficiency

Based on the findings from Study 1 about learners' sensitivity to context, Study 2 further explored how the ability to utilize contextual information for vocabulary learning relates to learners' proficiency in the target language. In particular, Experiment 5 found an

aptitude-treatment interaction (ATI) between the contextualized learning of vocabulary and learner' lexical proficiency in an intentional, sentence-based, and translation-given learning situation; specifically, the context imageability effect is selective about learner proficiency. The results suggest that even in a translation-based learning condition, contextual information is important for high-intermediate EFL learners. In contrast, beginners and low-intermediate learners might depend on a given translation regardless of the given context. The results also indicate that these groups' gain in form-meaning knowledge was at the halfway level of mastery; this was quite natural because the participants learned the target words only once in a limited time. Therefore, as assumed in Revised Hierarchical Model, and as suggested in Webb (2007a), the learners' major source of knowledge were the Japanese translations presented in the list. In Prince (1996), the upper and lower proficiency groups' TOEFL scores were 480 (equivalent to CEFR level B2) and 397 (level B1), respectively; these results are quite consistent with the assumption mentioned above if we regard Prince's upper and lower groups to be equivalent to high- and low-intermediate learners. However, it is difficult to compare the results of Prince's experiment and the current study because the learning modes were different: i.e., guessing-based contextualized learning and translation-based contextualized learning, respectively. Nevertheless, L2 learners at the high-intermediate level seem to be more sensitive to contextual information given during intentional vocabulary learning.

In the Common European Framework of Reference for Languages (CEFR; Council of Europe, 2001), levels A2, B1, B2, and C1 are labeled as follows: Waystage [Basic User 2] (A2), Threshold [Independent User 1] (B1), Vantage [Independent User 2] (B2), and Effective Operational Proficiency [Proficient User 1] (C1). Although a replicated framework for Japanese EFL learners called CEFR-J (Tono, 2012; see also Tono, 2013) and another framework developed based on a mega-study, called the EIKEN Can-do List (STEP, 2007; see also Dunlea, 2009) do not have such a set of labels, the CEFR labeling seems applicable to the Japanese situation. Unfortunately, neither CEFR, CEFR-J, nor the EIKEN Can-do List has

direct descriptors with respect to lexical skills or vocabulary knowledge because all of them are skill-based (i.e., listening, speaking, reading, and writing). However, can-do descriptors of learners' reading skills are worth mentioning in the current discussion. For reading ability, learners at the A2 level can "understand short narratives and biographies written in simple words" and "understand the main points of texts dealing with everyday topics" (e.g., Tono, 2013, pp. 294–295). This corresponds to the descriptions for Grades Pre-2 and 3 in the EIKEN Can-do List (STEP, 2007, pp. 12–15). In a similar manner of comparison, descriptors for the B1 level seem to correspond to Grade 2 (e.g., understanding the main points of English newspaper articles) and those for the B2 level to Grade Pre-1 (e.g., understanding texts dealing with topics of general interest such as current affairs). The descriptors for both the B2 level and Grade Pre-1 level suggest learners can adapt their reading style according to the text type and reading purpose. In Carver's (1992) terminology, learners achieve the ability to change the "gears" of reading at this proficiency level.

In the present study, the participants can be roughly grouped into four proficiency groups based on their scores on the first subsection of the STEP Eiken test. The first group is the beginner group, whose lexical proficiency was estimated at around Grade 3. As shown in Pilot Study 3, the participants at this level were able to understand almost all words in the simple contexts used in the present study (Experiments 3 and 5); however, they sometimes misinterpreted words and sentences. The second group, the low-intermediate group, was evaluated at Grade 2. These learners seldom misinterpreted the local meaning of the presented context. Although the present study cannot afford to provide detailed information about learners' can-do statements, one might speculate based on the description in CEFR-J that learners at this level were not able to switch reading modes according to their purpose or the given text type. This ability can also be regarded as the flexible use of "gears" (Carver, 1992; see Section 2.3.2). In contrast, this kind of flexible reading was expected for the third group, the high-intermediate group. In fact, the results from Experiment 6 focusing on

high-intermediate learners showed evidence of their flexibility in cognitive processes in context reading according to the text type (i.e., more and less imageable contexts). Although the different reading processes for more and less imageable contexts cannot be attributed to the skill of switching reading modes according to reading purposes, it seems highly consistent that the context effect was observed only among the learner group that was expected to have a more fluent command of reading modes than the less proficient groups. The high-intermediate group's lexical proficiency was higher than the low-intermediate group; however, their scores on the lexical proficiency test were below the pass line of Grade Pre-1. Finally, the participants who were expected to pass Grade Pre-1 are grouped into the fourth, or advanced, learner group.

In terms of the ATI between context imageability and learner proficiency, the beginner and low-intermediate groups did not show a clear effect of context imageability, implying that during intentional learning, their cognitive resources were devoted to the word association of L1 and L2 pairs or to the lower level processing of context (e.g., lexical and syntactic decoding; Grabe, 2009). On the other hand, the high-intermediate group showed a clear effect of context imageability in the test scores in Experiment 5. Given the findings from Experiment 6, it seems the learners at this level can achieve richer mental representations from the more imageable contexts than from those with less imageability. Taken together, it was quite natural to assume that the construction of rich mental representations or mental imagery enhanced the learning of new words in relation to contexts. In Dual Coding Theory, dual activation of imagery- and language-based processing is assumed to induce word learning (Paivio, 1991). In addition, it is assumed that the imagery-based global processing is usually faster than more analytic processing (see Kadota, 2006). Based on the different results for learner groups, it is be advisable for teachers to know that learners at different proficiency levels have different degrees of sensitivity to context. When students have developed their lexical proficiency to a certain level, they might be able to compare a new word's meaning

with the situation described in context. In such a case, their vocabulary knowledge gain will be maximized when a good context is provided. On the other hand, poorer learners cannot take advantage of mental representations. As mentioned earlier, the low-intermediate learner group in this study seems to correspond to the B1 level in CEFR. As long as Japanese learners are concerned, the B1 level is most related to high school students and lower level undergraduates (Tono, 2013, p. 171). This seems also compatible with Hasegawa's (2013) finding that the effect of context presentation on the impression of new words, which was observed among university learners, was not significant among high school students (see Section 2.4.3).

However, the context imageability effect seemed to disappear among the advanced learners. This phenomenon might be confusing because the advanced learners should be able to read most contexts without serious misinterpretation, apply proper reading modes according to their purpose and the text type, and construct rich mental representations from imageable texts. There seem to be at least three possibilities concerning the lack of context imageability effects among the advanced learners; this study briefly states these possibilities and then focuses on the most plausible one among them. The first possibility is that the advanced learners paid more attention to word association (i.e., L2 word forms and L1 translations) in the given word list than contexts. In other words, similar to beginners and learners in the low-intermediate group, learners at this stage were not sensitive to context type. However, this possibility does not seem quite plausible because the advanced learners should be able to read contexts fluently, without as much difficulty as the high-intermediate learners. One argument in support of this possibility might be that advanced learners thought the information about the word pairs was more important than contextual information. For example, they might have paid closer attention to the spellings of the target word compared to the high-intermediate learners. This study does not have data to examine this possibility; however, in the experiment, the participants were always instructed before the learning phase

that they should pay careful attention to word forms, meanings, and contexts. Thus, if the advanced learners followed these instructions, they must have read the contexts carefully at least once.

The second possibility was that the more advanced the learners were, the greater impact individual differences had. In other words, there might have been a greater variety of learner types in this proficiency group compared to other groups. As shown in a number of previous studies, lexical proficiency has a close relationship with vocabulary learning strategies (e.g., Fan, 2003; Kojic-Sabo & Lightbown, 1999; Lawson & Hogben, 1996; Maeda et al., 2003). In particular, studies suggested that learners can use more elaborated strategies as they become more proficient. This second possibility assumes that the individual differences within the proficiency group were larger in the advanced group than in the high-intermediate one, which would have made the context imageability effect unclear in this group. As shown in Experiment 4, learner types in terms of the use of vocabulary learning strategies seemed to affect the learners' sensitivity to context type. However, this view was not very plausible because the *SD* values, which indicated deviations of test scores within each learning condition, were not so different between proficiency groups in Experiments 1 and 2. In Experiment 2, the *SD*s were even smaller among the upper proficiency group than the lower group. Although there might have been a somewhat greater variety of strategies among the advanced group, it is rather difficult to consider this as the main cause of the insignificance of the context imageability effect.

The third interpretation is that the advanced learners were able to understand the less imageable contexts as deeply and fluently as the more imageable contexts. The present study supposes that all the three factors mentioned here might have affected the results (i.e., closer focus on word pairs, greater individual differences in vocabulary learning strategies, and better understanding of less imageability contexts, among the advanced learners than the other groups) but the third one had the most considerable influence. However, the discussion might

be complicated concerning the third possibility because of the following two problems. First, since Experiment 6 focused on the cognitive processes of high-intermediate learners, it was difficult to consider vocabulary learning by other proficiency groups including the advanced group. Therefore, a few additional participants whose lexical proficiency levels were even higher than the high-intermediate group in Experiment 5 were asked to participate in a think-aloud experiment similar to Experiment 6. The results are reported in the next section (Section 6.1.3). Second, some of the research results make the discussion difficult to understand. In particular, in Experiment 2, the test scores of the upper (i.e., advanced) proficiency level group were higher for the less imageable contexts than the more imageable contexts when the imagery instruction was given (see Section 3.4.3); this tendency was not statistically significant but worth discussing. Section 6.2 further discusses some apparently confusing results of the experiments, including possible reasons that the imageable contexts seemed to have a positive effect on the advanced learners' test scores (see Section 6.2.2).

Sections 6.1.3 and 6.2.2 suggest that the advanced learners' context reading during intentional vocabulary learning should be more carefully investigated in future research. Based on the observations in this study, it seems the learners at this level sought any interpretation, even from less imageable contexts. With respect to advanced learners' intentional vocabulary learning, Elgort (2011) reported that learners can gain implicit knowledge of new words by deliberate memorization of word cards (see also Nation & Webb, 2011, pp. 31–32). Although he did not describe it in detail, the target words were paired with both definitions and example sentences. For example, a pseudoword *proster* was presented on a word card with the definition "the part of the body comprising the hip, buttock, and upper thigh" and example sentence *This set of exercises focuses on the <u>proster</u> area*. If the present study's assumption is true, the advanced learners in Elgort's experiment might have generated a mental imagery regardless of whether or not the given contexts described a concrete and imageable situation. It will be interesting to examine if implicit knowledge can also be

acquired even if the word card does not provide any contextual information.

6.1.3 Different Processes of Context Reading

The two issues summarized and discussed in Sections 6.1.1 and 6.1.2 were mainly based on the test scores obtained from Studies 1 and 2. These findings suggest that it is important to consider how easily learners can construct an elaborated representation in their minds from reading contexts if they hope to learn new words more effectively in context. Effectiveness in the current context refers to how well learners can memorize new words associated with example sentences. However, one may argue that such seemingly effective learning might be not quite efficient in terms of cognitive processes; the learning can be either economical or uneconomical depending on how much cognitive burden learners spend on the mental process of comprehending the context. The discussion is even more complicated considering the efficacy in this sense because two different views might exist with respect to the relationship between learners' cognitive effort and their outcome of learning. First, one might regard one type of learning as more efficient than another if it requires less effortful cognitive processes. However, vocabulary learning is more effective when the learning includes more elaborated, and sometimes effortful, cognitive processes. For example, Involvement Load Hypothesis assumes that the degree of engagement is a critical factor in L2 vocabulary learning, which consists of different three aspects: (a) *need*, which refers to the requirement of having linguistic knowledge to perform a task, (b) *search*, which is related to learner's attempt to seek certain kinds of information such as word meanings, and (c) *evaluation*, which refers to the learner's reconsideration of what he or she has thought or learned (Hulstijn & Laufer 2001; Laufer & Hulstijn, 2001). Schmitt (2010) explains the third factor (i.e., evaluation) as the comparison of information about a word with the context of use to see how the word fits in the specific context (p. 27). According to this hypothesis, context reading during intentional vocabulary learning might increase learner's evaluation, causing

the learning process to become more complicated and effortful at the same time.

To examine how and how much learners utilized contextual information when memorizing a list of new word forms, meanings, and contexts, Study 3 examined learners' cognitive processes on the basis of the think-aloud protocols, learning time data, and self-evaluations of imagery generation during learning. The results suggested that the participants' attention was mainly devoted to the word forms and translations even when contexts were provided; however, it was the context-based strategies that made a more notable difference for individual learners. In particular, the learners used more imageable contexts differently from less imageable contexts, suggesting that their mental process of context reading was more analytic in nature when less imageable contexts were provided. In contrast, when they were given more imageable contexts, their process tended to include the generation of mental imagery. As to the degree of mental effort devoted to the contexts, it was found that less imageable contexts required greater effort from learners.

Therefore, based on the second major finding described in Section 6.1.2, namely, that imageable contexts enhanced vocabulary learning among high-intermediate EFL learners, it was suggested that students spend more time learning less imageable items but learn more imageable items better. In other words, although the more imageable contexts enabled learners to call up a mental image for the vocabulary item being learned, such an elaborated process of imagery generation was not cognitively demanding for high-intermediate learners. On the other hand, the participants performed analytic comprehension processes more frequently during the learning of less imageable contexts compared to more imageable contexts; this might have been a burden for the L2 learners. As mentioned in Section 5.4.4, this result was compatible with previous a finding that it might be inefficient for L2 learners to read a long passage and that a minimal length of context is the best (Laufer & Shmueli, 1997). Taken together, these findings imply that teachers should carefully consider what type of context to present to their students.

In summary, the L2 learners in this study read more and less imageable contexts quite differently. The most interesting suggestion might be that the effective use of contexts is related to inferential understanding and imagery generation during reading. With respect to the inferential understanding process as opposed to analytic decoding, it seems helpful to consider Kadota's (2006) model that was established based on Dual Coding Theory (Paivio, 1986, 1991) and other recent research findings. According to Paivio's model (1986, 1991), L2 learners' reading of contexts involves two mental processes: a language-based and an imagery-based process. One of the most basic assumptions of Dual Coding Theory is that the most effective learning occurs when the imagery-based (i.e., global) component is activated simultaneously with the language-based (i.e., analytic) component. If this assumption is applied to Kadota's model, it can be expected that the most effective learning (or understanding) will occur when the learning process involves both analytic and global features. However, because learners comprehend contexts with only a limited amount of working memory capacity, their cognitive attention seems to be oriented toward either analytic or global processes according to the imageability of a given context, to some extent.

However, the above assumption is established based only on the high-intermediate learners' think-aloud protocols. As discussed in Section 6.1.2, it is desirable to collect think-aloud data from learners at different proficiency levels in future studies. As the first step in this direction, two additional participants were asked to take part in a think-aloud session to examine how differently more skilled learners use contextual information during intentional vocabulary learning. The two participants were undergraduates majoring in linguistics, and neither of them had participated in any of the experiments by the author. However, their proficiency was not as high as native-like language users. In particular, the data of one participant were excluded from analysis because her test scores on the first subsection on the Eiken test were much lower than the other participant's scores. The other participant's lexical skills were estimated at a level that was similar to the upper group in Experiments 1 and 2. To

ensure a situation where the learner could read contexts without effort, learning materials were changed from a set consisting of rather difficult contexts, which was used in Experiments 3 and 6, to another consisting of much easier sentences, which was originally prepared for Experiment 5. The procedure was almost the same as in Experiment 6, including the practice session, preannouncement, and post-learning tasks (see Section 5.4.2.3). However, after they reported whether they had raised any imagery during reading contexts (i.e., mental imagery rating), the participants were also asked to write down what kind of imagery they had in mind. Therefore, there were three differences between this additional session and Experiment 6: (a) easier contexts, (b) higher learner proficiency, and (c) an additional imagery description task in the new session. No statistical analyses were performed because the sample size was too small to generalize the results.

The results showed the ratings about whether the contexts raised imagery during the reading context were higher for the less imageable contexts (M = 0.90, SD = 0.22) than the more imageable contexts (M = 0.70, SD = 0.45). In fact, almost all the contexts induced imagery: Among ten contexts, seven raised a mental image and two raised a partial image that was not as elaborated as the ones for the first seven. The imagery description report showed that inferential understanding of the content of contexts occurred during reading both more and less imageable contexts. Examples of the more imageable contexts were (a) *Mary and Paul had a repast together in the Italian restaurant*, (b) *Kate is a new doctor at the sanatorium in London*, and (c) *All the children played in the narrow esplanade last Sunday*. The reported imagery for each of these contexts was, respectively, (a') "These two are in love. [I imagined] spaghetti Napolitana," (b') "[I imagined] a white building and a woman in a white frock," and (c') "[I imagined] a riverside road [in] the evening, a slope [covered] with grass, and children's voices." The original protocols were written in Japanese; the author translated them into English, adding the expressions in square brackets.

On the other hand, examples of the less imageable contexts were (d) *Then they asked*

me, "Why don't you indite another one?" (e) *They danced in the gloaming and talked with each other*, and (f) *They wanted to mosey over to you, but they did not*. These contexts seem not to specify a concrete situation. However, the participants raised imagery such as (d') "The *one* refers to an academic article. The writer [of this sentence] must be a student," (e') "[I imagined] a European-style stone pavement and a man and a woman laughing," and (f') "The *you* refers to an old man. [He has a lot of] spare time. [I also imagined] a townscape in Paris and a café." Interestingly, the participants interpreted the less imageable contexts by generating inferences such as guessing about the referents of proper nouns. Similarly, in Pilot Study 4, think-aloud protocols from three graduate students included some inferences from the less imageable contexts such as (g) *The rock star's simple lifestyle contrasted sharply with the extravagant habits of many others in his profession* and (h) *A: Darling, did you talk to Sylvia about the importance of attending all her classes at college? B: I did, but I felt a bit hypocritical. After all, I often skipped classes when I was a student*. Their verbal protocol included (g') "Right. That fits [my] image of a rock star" and (h') "This is, well, ... [She] says *darling*. This is [his] wife's comments, I guess. Sylvia might be [their] daughter."

Both the additional participant described earlier in this section and the graduate students in Pilot Study 4 were expected to have higher English proficiency than the participants in Experiment 6. Therefore, more skilled learners seem to interpret the less imageable context by making inferences about some unspecified things such as the referents of pronouns, the role of a character, the relationship between characters, and so on. Based on these observations, it was speculated that learners at an advanced level are likely to seek any interpretation, even from less imageable contexts. This might explain the lack of significance in the context imageability effect among the upper groups in Experiments 1, 2, and 4; however, a more careful investigation is necessary in future research.

As summarized earlier in Table 6.1, this study found that more imageable contexts may enhance vocabulary learning but that the context imageability effect is selective about learner

proficiency. Imageable contexts are better remembered in relation to new words. In addition, it was found that learners read more and less imageable contexts differently; in particular, the effective use of contexts is related to inferential understanding. Based on these major findings and further discussion in Section 6.1.3, this study suggests the possibility that there are different stages of learner proficiency in terms of how sensitive learners are to context imageability. First, beginner learners do not seem to utilize a given context regardless of the context imageability, because they are not able to read contexts fluently and strategically (Griffin, 1992; Prince, 1996). As suggested in Prince (1996) and Nassaji (2006), learners' effective use of contextual information should be closely related to their lexical ability (see Section 4.3). Without sufficient knowledge and skills of vocabulary and grammar, learners cannot achieve a propositional understanding of what is written in the contexts (see Section 2.3.1). However, contextual information might be more comprehensible for learners in language classrooms if teachers provide appropriate instructions such as an oral introduction to the reading material and explanation about its contexts.

Second, learners with a certain level of L2 proficiency can read the context with some degree of fluency during intentional vocabulary learning, which enables them to understand not only the literal meaning of the context but also more implicit information when the context is imageable for the learners. In other words, learners at this stage seem to be most sensitive to context quality. It is difficult to define what exact level and type of proficiency is necessary for this stage. However, based on Study 2 in this investigation, it seems reasonable to predict that the learners' lexical proficiency that corresponds to a level higher than the "Threshold" indicated in CEFR (Council of Europe, 2001; see Section 6.1.2). If learners are above this level, or if the contexts are quite easy for them to understand, both more and less imageable contexts might be equally effective in intentional vocabulary learning. According to the observation described in Section 6.1.3, advanced learners might be able to construct a rich mental representation by generating inferences about information that is not explicitly

stated in the less imageable contexts (e.g., referents of proper nouns).

6.2 Three Specific Problems and Possible Solutions

The present study aims to further summarize and organize these facts and findings to propose a working model that explains the role of context reading in the translation-based intentional vocabulary learning. However, there were a few confusing facts in this study; the three subsequent subsections address these issues. First, although it has been already discussed earlier, the effect of learner proficiency seemed to be reversed in Experiment 3, in that the upper level learners were more sensitive to context imageability. Section 6.2.1 briefly explains why this result was obtained. Second, the effect of context imageability seemed to be reversed in one condition in Experiment 2, although this tendency was not statistically significant. This problem was explained in Section 6.2.2. Third, although it was also an insignificant effect, one learner group seemed to be more sensitive than the other two groups in Experiment 4. The apparently sensitive group used both information about word pairs and contexts, whereas the other two groups used either word pairs or contexts. This issue is discussed in Section 6.2.3. The discussions in these three subsections support the possibility mentioned earlier in this section that there are different stages of learner proficiency in terms of how sensitive they are to context imageability.

6.2.1 Effect of Learner Proficiency in Experiment 3

To test whether translation-based learning of vocabulary can be affected by differences in the context in which target lexical items are presented, Study 1 included three experiments (Experiments 1, 2, and 3). The major finding from each experiment is concisely presented in Table 6.2; the present study obtained evidence in a step-by-step manner. Figure 6.1 displays the results of the contextualized recall test of the target words in the three experiments; however, in addition to Experiments 1, 2, and 3, the comparable results from Experiment 4

are illustrated together. As shown in Figure 6.1, the relationship between learner proficiency and context imageability seemed to be reversed in Experiment 3. Namely, the effect of learner proficiency was significant for the less imageable contexts in Experiments 1 and 2 but the more imageable contexts in Experiment 3. However, these seemingly confusing results should be interpreted from another viewpoint: learners' sensitivity to context imageability. From this point of view, the results can be explained as follows: The "lower" level groups in Experiments 1 and 2 as well as the "upper" group in Experiment 3 were more sensitive to the context type effect than the other group(s).

Table 6.2

Summary of the Major Findings in Experiments 1, 2, and 3

Experiment	Proficiency	Major finding
1	High intermediate	The lower proficiency group was sensitive to context quality.
2	High intermediate	A contextualized test of lexical proficiency should be used.
3	Beginner	The upper proficiency group was sensitive to context quality.

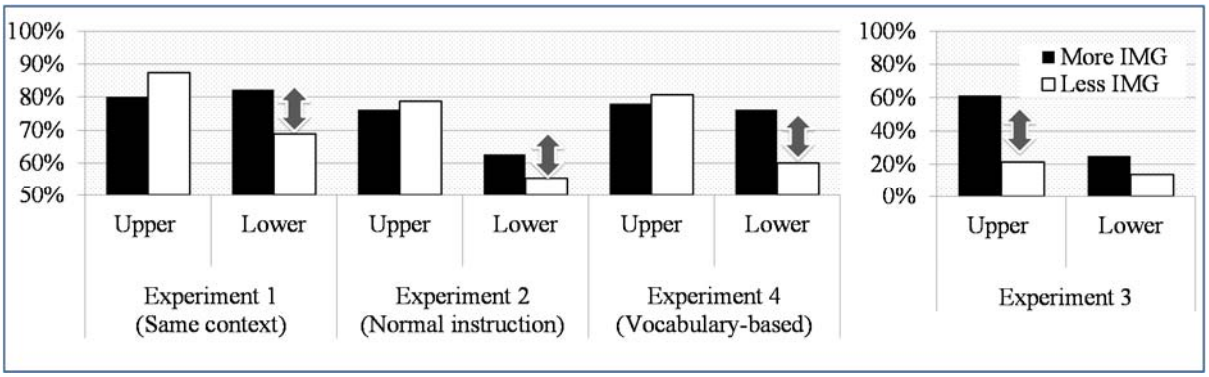

Figure 6.1. Correct answer rates (%) of contextualized tests in Experiments 1, 2, and 4 compared with those in Experiment 3. As the black arrows suggest, the positive imageability effect was notable among the lower proficiency groups in Experiments 1, 2, and 4, whereas it was outstanding among the upper proficiency group in Experiment 3.

In Experiment 1, Japanese university students learned a list of pairs of unfamiliar English words and their Japanese translations, with example sentences. To examine the sensitivity of these EFL learners to context information in this learning condition, two types of context were prepared: more and less imageable contexts. Although this study was not designed to compare contextualized and decontextualized learning, the results showed a kind of ATI between context type and learner proficiency: The test scores of the lower proficiency group were reduced when they were given lower imageability contexts. Also, Experiment 2, which initially focused on proficiency measures, found an ATI between context type and learner proficiency consistent with the results of Experiment 1. The results of Experiment 3 seem to be different from those of the earlier two studies, because the effect of context imageability was found in the upper proficiency group, not the lower group. This difference was considered to be caused by the difference in the overall proficiency level of participants between the studies; the participants in Experiments 1 and 2 were mostly high-intermediate learners, whereas those in Experiment 3 were beginners, with a few exceptions. In fact, the "lower" proficiency groups in Experiments 1 and 2 were actually more skilled learners than the "upper" group in Experiment 3. However, it was difficult to compare the learner groups directly across the studies because Experiment 3 did not conduct the same proficiency test, which seemed quite challenging for the beginners. In addition, the number of participants was small in Experiment 3, which makes it difficult to accurately define their proficiency level. Nevertheless, the results seem to indicate that there is a certain range of proficiency at which imageability affects vocabulary learning. Therefore, the relationship of learner proficiency and the sensitivity to context imageability was further investigated with a wider range of participants in Experiment 5.

As described in Section 4.5.3, Experiment 5 showed that the high intermediate learners were the most sensitive to context. Based on the results, one may argue that a learner will follow a process of lexical development featuring the following steps: (a) S/he is not able to

deal with contextual information during intentional vocabulary learning, (b) s/he is at her/his most sensitive to context quality, and (c) s/he can always learn new words effectively, regardless of context type. This account was quite consistent with the ATI between context type and learner group found in the other experiments (i.e., Experiments 1 to 4); the "upper" group in Experiment 3 and the "lower" group in Experiments 1 and 2 would all have been in the second stage.

Thus, as previewed in the introductory part in Section 6.2, the following assumption was established. In the first stage, beginner learners are not capable of taking advantage of contextual information regardless of the context imageability, as long as translation-based deliberate memorization is concerned. Second, learners at a certain level of L2 proficiency can read the context with some degree of fluency during intentional vocabulary learning; they can understand not only the literal meaning of the context but also more implicit information when the context is imageable. Third, if learners are above the high intermediate level, or if the contexts are quite easy for them to understand, both more and less imageable contexts might be equally effective in intentional vocabulary learning because learners might be able to construct a rich mental representation by generating inferences about information that is not explicitly stated in less imageable contexts (see Section 6.1.3).

6.2.2 Effect of Context Imageability in Experiment 2

The second problem concerns a kind of seemingly reversed effect of context imageability observed in one condition in Experiment 2. That is, the test scores of the upper group in Experiment 2 were higher for the less imageable contexts than the more imageable contexts when the participants were required to imagine the situation described in the contexts (see Section 3.4). Although this tendency was not statistically significant, some sort of explanation is necessary. As mentioned earlier, the upper group in Experiment 2 can be characterized in terms of their lexical proficiency, which was higher than that of the typical

high-intermediate learners (see Section 6.2.1). Although these participants could be regarded as either (an upper range of) high-intermediate learners or (a lower range of) advanced learners, in the present study, it was expected that their behavior would resemble advanced learners'. Based on CEFR (Council of Europe, 2001) and other CEFR-based frameworks of language proficiency (STEP, 2007; Tono, 2012; see also Dunlea, 2009; Tono, 2013), the upper group in Experiment 2 might have been able to adapt their reading style according to the text types and reading purposes (see Section 6.1.2).

The tendency that the upper group showed the imageability effect when they were given imagery instruction seem to be closely related to the fact that they did not show the imageability effect without any special instruction. Concerning this issue, Section 6.1.2 presented three possible reasons for the lack of context imageability effect among the advanced learners: (a) closer focus on word pairs, (b) greater individual differences in vocabulary learning strategies, and (c) better understanding of less imageability contexts. All these factors might have affected the results; however, based on the observations in this study, the third one seems the most plausible. In other words, learners at this level seem to seek plausible interpretations spontaneously even from less imageable contexts. Furthermore, as mentioned in Section 6.1.3, when learners can understand the context considerably easily, they can make inferences to elaborate their comprehension of it. Especially, when the context written in plain English describes an unspecified situation (i.e., the less imageable context), it seems that learners tend to generate inferences so that the context makes sense. For example, when a participant in the present study read the less imageable context *"Why don't you indite another one?"* he reported that he thought *one* referred to an academic article and the writer of this sentence must be a student (see Section 6.1.3). He also reported that the context evoked mental imagery while he was studying it.

However, it is important to point out that learners with relatively high proficiency tended to attempt to interpret abstract contexts by generating inferences in the think-aloud

task. In other words, the frequent generation of such inferences was likely induced by the experimenter's instructions for the participants to report everything they came up with during learning (i.e., the think-aloud method). Therefore, it might be because they were told to focus on the meaning of the context that they reported imagery based on both a literal and inferential understanding of the contexts. If this assumption is true, one may predict that other instructions would enhance the more skilled learners' inferential understanding of the contexts. The imagery instruction, given to the participants in Experiment 2, might have functioned as a task that forced the participants to deeply understand the contexts, at least for the upper proficiency group. In the present study, the instructions involved two components: (a) a request for the participants to formulate a mental image of the text and (b) rating the imageability of each context, which ensured that all participants properly considered the content of the contexts (see Section 3.4.2.3). As introduced in Section 2.4.1, past studies on reading comprehension used the imagery instruction when the researchers intended to help readers construct highly elaborated representations (e.g., Alba, 1984; Gambrell & Jawitz, 1993; Horiba, 2002). Although it was a study on children's reading in L1, Gambrell and Jawitz (1993) found an effect of imagery instruction on text comprehension. Because the participants were fourth-grade students and not quite skilled readers, they were not able to activate the imagery system assumed in Dual Coding Theory during reading. Gambrell and Jawitz assumed that illustrations acted as a bridge or transitory step in the process of transforming written information into mental images.

In the present study, it seemed difficult for the L2 learners to activate the imagery system when they read the less imageable contexts. Based on this assumption, the imagery instruction might have functioned as an enticement to transform the verbatim understanding into more imaginative comprehension based on inferences. Research has shown that inference generation plays an important role not only in L1 but also L2 reading; however, the inferential comprehension of contexts is influenced by L2 proficiency (Horiba, 1996; Ushiro et al., 2011,

2012; Yoshida, 2003). Therefore, the lower proficiency group in Experiment 2 was not able to generate inferences to draw interpretations from the less imageable contexts. The discussion in this section is a possible explanation of the problem that when the advanced learners were given the imagery instruction, the effects of context imageability seemed to be reversed. However, this tendency was not statistically significant, and therefore, more careful investigation is necessary in future research.

6.2.3 Differences Between Learner Groups in Experiment 4

The third problem that might have caused the readers' confusion was the tendency found in Experiment 4 that one learner group seemed to be more sensitive than the other two groups, although this effect was not statistically significant (see Section 4.3). In the analysis based on learner groups, three types of learners were identified: (a) word-oriented, (b) context-oriented, and (c) word-and-context-oriented. Compared to the first two groups, the third group, which used both information about word pairs and contexts, seemed to be the most sensitive to context imageability. This tendency was interesting but rather confusing because the context-oriented learners were not as sensitive as the word-and-context-oriented ones. This section discusses why the context imageability factor seemed more important for the word-and-context group.

As mentioned earlier, when readers devote their attention to memorizing new words, they are likely to comprehend the context using only limited cognitive resources (see Sections 2.3.2 and 5.4.3). In fact, Experiment 6 showed that around two thirds of the cognitive strategies used during intentional vocabulary learning were word-based processes. With regard to the learners' context reading, most of them comprehended each context carefully and often repeatedly. However, their reading comprehension tended to be superficial for the less imageable contexts, partially because their predominant goal was to memorize the target words (i.e., word forms, meanings, associations of the word forms and their meanings, and

word use in a context). One may suppose that if the learners were to comprehend the example sentences for other purposes such as understanding the writer's intention, imaging the situation described in the contexts and memorizing the content of the contexts, then they might be able to focus on the contextual information more closely. Considering this, one possible explanation for the word-and-context group's sensitivity to context imageability compared to the context-oriented group was that their cognitive resources devoted to the context-based processes were relatively limited. It is worth noting that Experiment 2 was conducted with a time limit. Therefore, the word-and-context-oriented group probably hurried to learn both the word-based and context-based information; on the other hand, the context-oriented group should had more time to comprehend the contexts by concentrating on the context-based strategies.

Another explanation for why the word-and-context group showed more sensitivity to context imageability than the other groups was that the mental process of associating the information of the target words and contexts was the most obvious in this learner group. As mentioned earlier (see Sections 2.4.3, 3.2.4. and 6.1.1), reading a context embedding a new word can affect the readers' recognition of the new word, resulting in the construction of a mental association between the word and context (i.e., word-context association). As suggested in some theoretical models, there are gaps between translation pairs between two languages (e.g., *water* and *clutch* in English paired respectively with *mizu* and *nigiru* in Japanese); to fill in the gaps and improve the learners' L2 mental lexicon, repeated contextual input is necessary. It is assumed that incremental development of vocabulary knowledge in this sense can be realized by experiencing a number of opportunities to encounter different contexts, as discussed in Section 6.1.1. Thus, it should be necessary to construct word-context associations or elaborate lexical knowledge by comparing a word to its context information each time it is encountered in a context to achieve the incremental learning of unfamiliar words. According to the findings and suggestions in the present study, the ease of

construction of word-context associations was influenced by context imageability. The current account of the word-and-context group's learning regards that this group's test scores were affected more greatly by context imageability than the other two groups' scores (i.e., the word-oriented and context-oriented groups) because the learners in this group tended to exert effort to learn the relationship between the target words and their contexts.

However, there is still another account, which can directly refute this explanation. Namely, the tendency was caused by a kind of *spurious relationship* between the learner type and the sensitivity to context imageability; the present study cannot reject this possibility, and therefore, Section 4.3.4 avoided making a conclusion with respect to the learner group effect. In general, the term "spurious relationship" refers to a situation where two variables appear to have any causal relationship, but actually, the relationship is due to either coincidence or the presence of a certain third factor that causally relates to both the first and second variables. In the current case, the third variable might be the participants' lexical (or general) proficiency in English. Considering that the repetition of word forms and translation is the most basic vocabulary learning strategy among Japanese learners (e.g., Maeda et al., 2003), simultaneous use of both word forms, translations, and contexts might be an effortful process for the learners. Therefore, it was plausible that the word-and-context-oriented group included more skillful learners. The current account assumes that it was the learners' lexical proficiency that causally related to their sensitivity to context imageability.

Unfortunately, past studies do not provide sufficient evidence to determine which of the three possible accounts is closer to the truth: the word-and-context group's (a) limited cognitive resources, (b) increased efforts to learn the relationship between the target words and their contexts, or (c) higher lexical ability or other factors. In addition, it is risky to make a conclusion on this issue because the relationship between the learner group and context imageability was not statistically significant. Therefore, this problem should be further investigated in future research. One possible experimental design might be to compare word-,

context-, and word-and-context-oriented learner groups with similar English proficiency levels; conducting an intentional learning experiment using both more and less imageable contexts, the possibility of the third account (c) can be more closely examined. In addition, to differentiate the first (a) and the second (b) possibilities, a similar experiment comparing two conditions with and without a time limit will be also suggestive.

6.3 A Working Model Focusing on Mental Imagery Generation

Through the six experimental studies as well as the four pilot studies, the present study collected various data on Japanese EFL learners' intentional vocabulary learning using a list of English word forms, Japanese translations, and English contexts. Although there are some unsolved issues that should be further investigated in the future (see Section 6.2), the suggestions and findings summarized and discussed in Section 6.1 seem to be informative for both researchers and practitioners of English education in Japan. However, there were a number of variables related to each other in this study, such as the presence of contexts, imageability and other features of contexts, English lexical proficiency and other learner factors, imagery instruction, preannouncement and types of vocabulary tests after learning, and learning modes (e.g., intentional vs. incidental) and reading purposes (e.g., reading for learning vs. rauding); therefore, one may argue that the implications should be further summarized and organized. To present an overview of the findings in this study, this section proposes a working model that explains how L2 learners use the word-based and context-based information presented in a list during intentional vocabulary learning. The most important assumption in this model is that imagery activation during context reading may affect the efficacy of contextualized learning.

First, this study's focus is illustrated in Figure 6.2. Although the findings might be also suggestive for incidental vocabulary learning (see Section 2.2.2), this study focused on intentional vocabulary learning using a word list consisting of the target word forms in L2,

their typical meanings in L1, and contexts in L2. The general instruction was always given to learners in Experiments 1 to 6, asking them to learn these three things as much as possible. Some of the external factors illustrated in Figure 6.2 were different across the experiments; however, the most typical learning mode in this study was one with (a) preannouncement of tests after learning and (b) time limit that requires participants to learn the word list speedily. These two factors ensure that the participants were engaged in intentional vocabulary learning, rather than any other modes of vocabulary learning (e.g., Schmitt, 2010). As for the teacher instruction, the current section regards the normal instruction condition as the most typical learning situation.

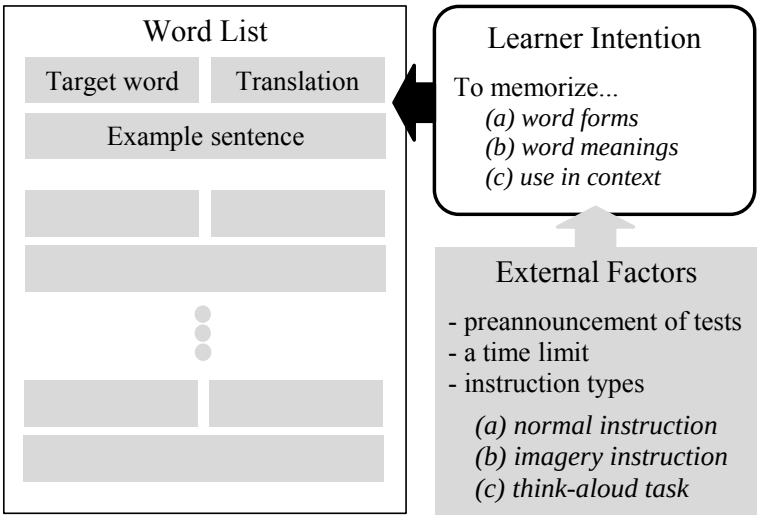

Figure 6.2. The learning mode focused on in this study.

In this mode of vocabulary learning, learners pay attention to word forms and their translations because the first stage of learning a new word is often word association between L2 and L1. This assumption is common in psycholinguistic studies that support Revised Hierarchical Model (e.g., Jiang, 2000; Kroll & Stewart, 1994; Kroll et al., 2010; Sunderman & Kroll, 2006) and other experimental studies focusing on vocabulary learning (e.g., Griffin, 1992; Laufer & Shmueli, 1997; Prince, 1996; Webb, 2007a) and literature that overviews

research findings (e.g., Folse, 2004; Nation, 2005, 2009, 2013; Nation & Webb, 2011; Schmitt, 2010). According to Webb (2007a) and other related studies, the major learning source in the learning list is the word forms and their translations; context has a smaller effect as a knowledge source when the word meanings are explicitly presented in L1. In the present study, the cognitive strategies using the target words themselves and their translations are called word-based strategies, whereas those using example sentences are context-based strategies. As illustrated in Figure 6.3, the word association process is directly related to the word-based strategies. On the other hand, as discussed in Section 6.1.1, the context-based strategies seemed to result in the mental association between the lexical information and contextual information. This kind of connection constructed in learners' minds can be called word-context association. Supporting the prediction from prior research findings, this study found that the frequency of word-based strategies was greater than context-based strategies, suggesting that word association was more important than word-context association for learners memorizing new words; in Figure 6.3, the size of the boxes for these two types of mental association reflects the relative importance. Also, recent studies suggested that L2 learner's attention to word forms, which can be reflected in the time they spend looking at each word (i.e., eye-fixation time during reading a text; Godfroid et al., 2013).

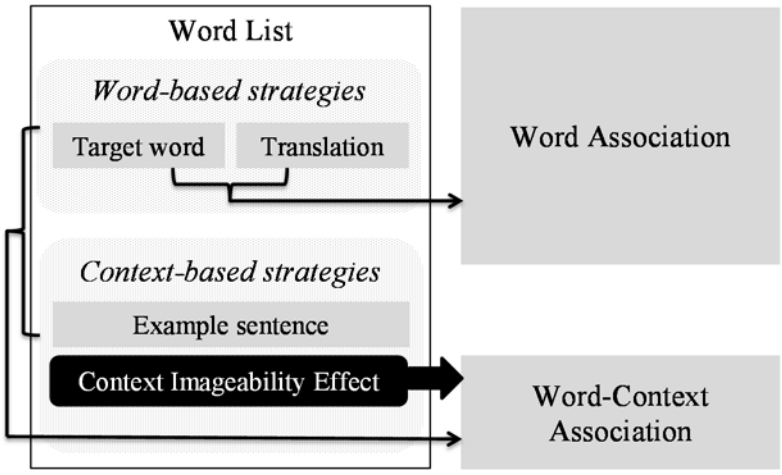

Figure 6.3. The two types of cognitive strategies and assumed mental associations.

Because of the relatively small role of word-context association in intentional vocabulary learning, one might suppose that there is no need to present contexts to learners. In contrast to the presumption that L2 learners would be likely to ignore contextual information, this study repeatedly found that learners can actually be sensitive to the content of the contexts. This finding was obtained by experiments that took into consideration the ease of imaging the situation described in the context (i.e., context imageability). As described in Figure 6.3, it can be assumed that context imageability affects the learners' rates of constructing word-context associations in their mind. According to Dual Coding Theory (Paivio, 1986, 1991, 2006), items in a list (e.g., real words or pseudowords) are learned better when they are associated with imageable items (e.g., concrete words) in comparison to less imageable items (e.g., abstract nouns) because learners' memory of the association is enhanced by the imagery activation. This assumption is consistent with Hasegawa's (2010, 2013) results, as well as our everyday experience that we can remember things well when we have a strong impression or vivid image related to the things to be remembered.

Furthermore, the results in this study suggested that more imageable contexts can enable learners to remember the target words in a cued recall test (Study 1) because the imageable contexts might lead to more effective cognitive processes in context reading (Study 3). The relationship between the context imageability, typical reading processes, and the learning outcomes assumed in this study is displayed in Figure 6.4. According to the results of the think-aloud study (Experiment 6), more imageable contexts were more likely to be comprehended with in a shorter time, accompanied with imagery generation. This finding is compatible with Begg and Paivio's (1969) suggestion that when learners read an imageable sentence, they tend to retain the overall meaning; on the other hand, when they read a less imageable sentence, they tend to remember the superficial wordings. Presumably, the process of mental imagery generation during vocabulary learning is closely related to the construction of the word-context associations in mind. According to Kadota's (2006) model, L2 learners'

reading of contexts involves two mental processes: analytic and global processes. These two processes seem to relate with the two components assumed in Paivio's (1986, 1991) Dual Coding Theory. One of the most basic assumptions of Dual Coding Theory is that the most effective learning occurs when the imagery-based (i.e., global) component is activated simultaneously with the language-based (i.e., analytic) component. If this assumption can be applied to L2 readers' comprehension processes, it can be expected that the most effective learning (or understanding) will occur when the learning process involves both analytic and global features.

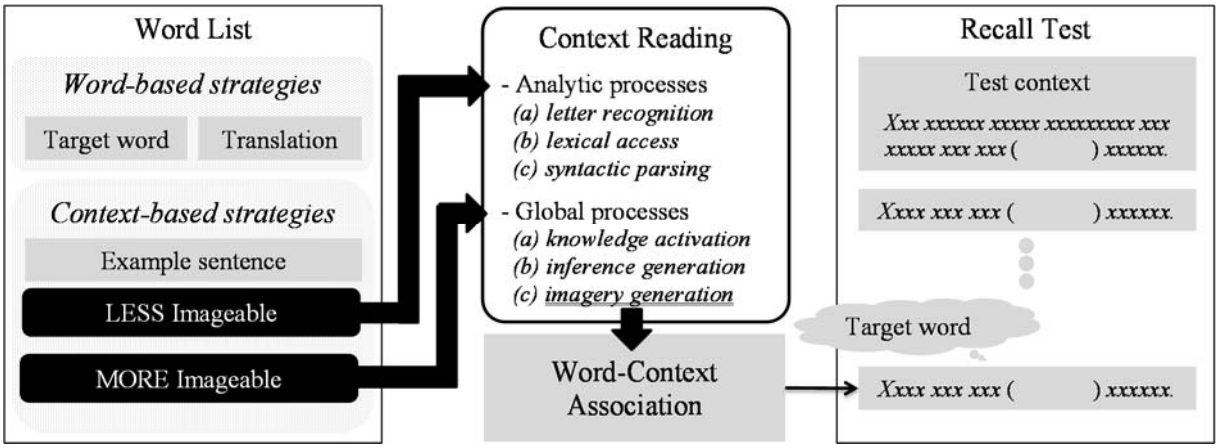

Figure 6.4. Effects of context imageability on reading processes and recall test scores.

To improve Figure 6.4 in terms of its elaborateness and informativeness, one might suppose that previous theoretical models focusing on L2 learners language comprehension such as Paivio's (2006) Bilingual Dual Coding Theory and Kadota's (2006) model can be applied to the present study's framework as a working model of the mental processes in contextualized vocabulary learning. A new model might illustrate how the vocabulary learning processes relate to the learner's long-term memory, including knowledge of newly learned lexical items and memory of recently read contexts. The internal structure of the context-based processes should be illustrated considering the language-based or analytic

processor and the imagery-based or global processor proposed in the original models by Paivio and Kadota. However, the present study does not present a specific figure describing such a revised version because there have been no evidence about the analytic and global processors and the strategy controller should work to support the learner's word-based and context-based processes.

Nevertheless, the current model seems quite useful for the discussion of the present study's data because it offers further suggestions about how the imageability factor affected the learning processes. The model reflects three assumptions. First, it assumes that learners' cognitive attention during intentional vocabulary learning is mainly devoted to word forms and translations. Second, the model suggests that learners comprehend contexts with only a limited amount of working memory capacity. To optimize working memory performance, the learner's cognitive attention is oriented toward either analytic or global processes according to the imageability of a given context. Third, the model assumes that when the global processor generates mental imagery from a context, this imagery can be associated with the memory of target words in some way. Based on this assumption, when learners are given more imageable contexts for the learning of unfamiliar words, they should be able to recall the combination of target word and contexts. In fact, Experiment 5 found that university students at a high-intermediate English proficiency level more target words when imageable contexts were given as cues; in addition, Pilot Study 1 showed that when given target words as cues, they recalled the content of more imageable contexts better than that of less imageable contexts. The current model's assumptions are quite consistent with these results.

However, the vocabulary learning processes described in Figures 6.2, 6.3, and 6.4 were based on the high-intermediate learners. Study 2 discovered that the context imageability effect was unique to learners with that level of lexical proficiency. In contrast to high-intermediate EFL learners, beginners and low-intermediate learners might depend on a given translation regardless of the given context because the participants learned the target

words only once in a limited time. As Webb (2007a) suggested, the learners' major source of knowledge was Japanese translations presented in the list. On the other hand, more advanced learners than the high-intermediate leaners might also learn new words similarly between more and less imageable context conditions. As Section 6.1.2 has proposed, advanced learners might (a) focus on the word pairs closer to prepare for a posttest, (b) have greater individual differences in vocabulary learning strategies, and (c) comprehend less imageability contexts more easily, compared to the intermediate learners. In particular, based on the third possibility, learners at this level seem to seek plausible interpretation spontaneously even from less imageable contexts (see Section 6.1.2). Moreover, when the context written in plain English describes an unspecified situation, it seems that the advanced learners might generate inferences so that the context makes sense (see Section 6.1.3). Further investigation on cognitive strategy use considering learner proficiency is necessary to improve the working model proposed in this study.

Chapter 7

Conclusion

7.1 Status of the Current Thesis

This study focused on intentional vocabulary learning using a list of new L2 words paired with their translations and example sentences and explored the role of context reading. According to past studies, contexts seem to be unnecessary in translation-based learning, as opposed to guessing-based learning. Researchers have introduced only a limited amount of evidence showing in what conditions L2 learners can take advantage of contextual input during vocabulary learning. However, this study discovered the effect of context on the scores of a vocabulary test administered after the learning of a word list by taking into consideration two important factors: the context imageability and learner proficiency. The major findings of the six experiments conducted in this study are summarized in Table 7.1, which is the same as Table 6.1 in Section 6.1. As shown in the table, this study has two main focuses: (a) how teachers and learners can optimize the effect of context reading for vocabulary learning (i.e., condition) and (b) how learners comprehend contexts while their intention is oriented toward memorizing lexical items (i.e., process). The research hypothesis was that, after learners encounter lexical items in a context, they retain the contextual information in relation to the target lexical items. Some learners retain context by drawing a picture in their minds of the situation described. Although this section does not repeat the detailed description of the results, it should be emphasized that this study provides evidence to promote the contextualized learning of vocabulary. Interestingly, as far as the contexts were simple and imageable, context presentation did not reduce learning efficacy among any learner groups.

Table 7.1

Summary of the Six Experiments, Three Viewpoints, and Two Focuses in This Study

Experiment	Viewpoint	Focus 1: Condition	Focus 2: Process
Experiments 1, 2, and 3 (Study 1)	(a) sensitivity to contextual information	More imageable contexts may enhance vocabulary learning.	Imageable contexts are better remembered in relation to new words.
Experiments 4 and 5 (Study 2)	(b) interaction between imageability and proficiency	The context imageability effect is selective about learner proficiency.	Poorer learners cannot take advantage of mental representations.
Experiment 6 (Study 3)	(c) different processes of context reading	Effective use of contexts is related to inferential understanding.	Learners read more and less imageable contexts differently.

Note. This table is the same as Table 6.1 in Section 6.1. Experiments 1, 2, and 3 constituted Study 1, which explored the possibility that L2 learners' encoding (i.e., learning) and retrieving (i.e., recall) processes are affected by the context type. Experiments 4 and 5 constituted Study 2, which further examined the nature of the ATI (i.e., the interaction between the learner group and context type) found in Study 1. Experiment 6 and additional data related to it were discussed within the framework proposed in Study 3, which focused on the cognitive strategies during learning a word list with example sentences.

Based on the empirical findings from Studies 1 to 3 and theoretical frameworks established in the previous research, Section 6.3 proposed working models to illustrate the overview of the relationship between major factors and processes considered in this research. However, the figures displayed to explain the models might give readers an impression that reading context is not quite important after all, since the role of contexts does not seem as substantial as that of translations, as far as the intentional learning of a word list within a

limited time is concerned. Although researchers now seem to agree that the effective use of translation is a more economical approach to learning new words than guessing-based learning, the role of context should not be underestimated. As Nation (2009, 2013) pointed out, the most effective method of L2 vocabulary learning is a combination of intentional and incidental modes of learning. Specifically, he recommended that teachers should design an L2 class considering the balance of the following four activities: (a) meaning-focused input, (b) meaning-focused output, (c) language-focused learning, and (d) fluency development (i.e., the *four strands*, in his terminology). In the Japanese educational context, it is suggested that after learners first memorized the association of an unfamiliar word in intentional learning, they should encounter that word repeatedly in various contexts to enrich and modify their previous knowledge about it (e.g., Tono, 2013).

Considering the combined approaches of intentional and incidental vocabulary learning, the model that was introduced as Figure 6.4 can be arranged as in Figure 7.1 below:

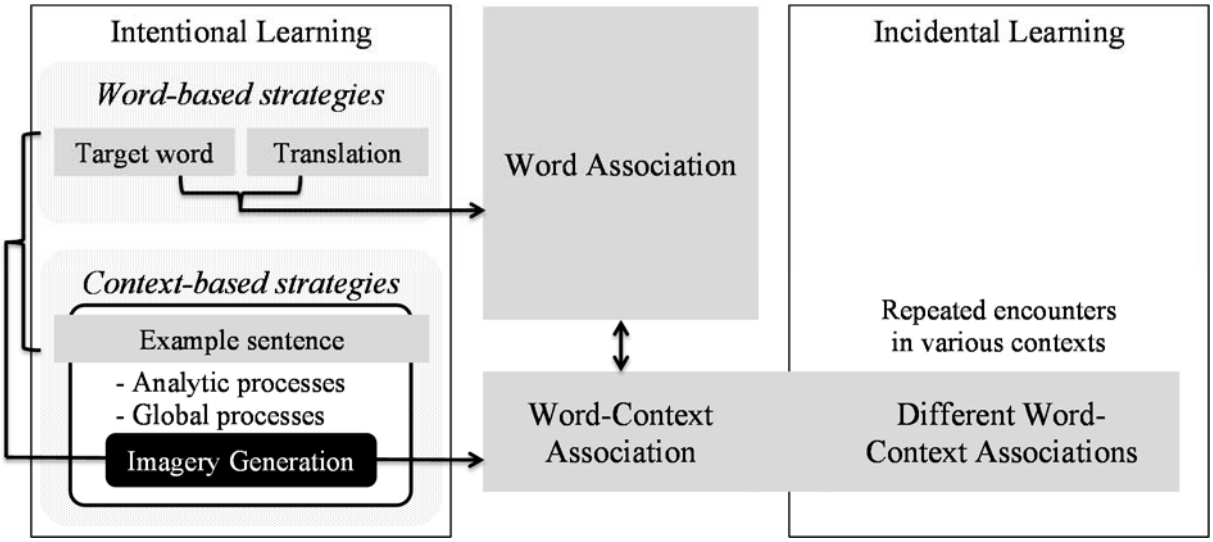

Figure 7.1. The assumed function of reading context in intentional learning. It is also illustrated that learners may experience further contextual inputs.

The two largest boxes represent intentional and subsequent incidental learning. When

learners first encounter a new word during intentional vocabulary learning, their attention is devoted to the word form and meaning, resulting in a construction of an L2-L1 word association in their minds. If contexts are also provided, the learners might read them through analytic and global processes; the balance of these two processes is affected by the ease of generating mental imagery from the context (i.e., the context imageability effect). This phase seems to determine the degree of constructing the word-context association in one's mind. This is why Figure 7.1 shows an arrow from the word- and context-based processes to word-context association via imagery generation. Imagery generation is placed just below the global processes of context comprehension because the two are closely related.

Based on the assumptions that learners use contextual information to evaluate what they have understood about the target words, this model regards that learners' knowledge concerning word and word-context associations may interact in some way. For example, some participants in Experiment 6 produced new translations of the target words that matched the content of the contexts. Figure 7.1 is different from the previous figures because it assumes that learners encounter a new word repeatedly in different contexts through message-based input and output and fluency development procedures (Nation, 2009, 2013). Through incidental learning experiences with different contexts, learners enrich their conceptual representation of the new word. The basic assumptions concerning L2 learners' mental lexicon have already been introduced in Chapter 2.

As in Chapter 1, suppose a learner reads a context such as *Mrs. Smith was a very <u>pious</u> woman who attended church services regularly* and he or she is not familiar with the word *pious*. The learner might retain *pious* in relation to an image aroused by the context concerning the described person, *Mrs. Smith*. If this is the case, then learners can recall the meaning of the target word *pious* in relation to the context and vice versa. If this learning accompanies an L1 translation such as *shinjin-bukai* in Japanese, the word association between the word pair (i.e., *pious* and *shinjin-bukai*) could be reinforced by the activation of

imagery from the context, as assumed in Dual Coding Theory (Paivio, 1986, 1991). Also, the presence of context might be helpful for learners to bridge the gap between the concepts of *pious* and *shinjin-bukai*, as assumed in the Distributed Conceptual Feature Model (van Hell & de Groot, 1998; see also de Groot, 1992; Kroll & de Groot, 1997). However, learners' lexical knowledge cannot be the ideal state, which is proposed in Jiang's (2000) theory based on the Revised Hierarchical Model (Kroll & Stewart, 1994), until they experience a sufficient amount of additional contextual input. For example, they need to know that the term *pious* can be used in many ways such as *pious acts*, *pious sentiments*, and *pious hopes*. It is because there are too many aspects to be learned about this word that incidental learning is important even if the learners have already understood one meaning of the newly learned word (see Nation, 2009, 2013; Tono, 2013). Nevertheless, implications gained from the present investigation seem to be informative in considering the function of context reading during translation-based learning of a word list.

7.2 Limitations and Suggestions for Future Research

Although the current research may give informative suggestions for L2 vocabulary instruction, some limitations need to be recognized. In particular, there are three important points that should be further investigated in future research; they are concerned with the three specific problems discussed in Section 6.2. The first problem, mentioned in Section 6.2.1, is the complicated interaction of the learner proficiency factor and context imageability. In particular, the relationship between L2 learners' sensitivity to context types and their lexical proficiency should be more carefully examined. Unfortunately, the specific items of the proficiency tests used in the present research were not always identical across the experiments. Therefore, the proficiency test scores could not be compared directly between the experiments. A replication study should be conducted using an experimental design that is similar to Experiment 5 in this study. However, based on the current research results, it seems reliable

that learners at a certain proficiency level are more sensitive to the context type than others are. Also, it might be interesting to investigate which type of lexical ability is the most closely related to context sensitivity during intentional vocabulary learning, since lexical proficiency can be assessed in many ways such as by vocabulary size, depth, and fluency.

Second, as explained in Section 6.2.2, some learners who were regarded as the most proficient participants in the present study showed an interesting tendency in terms of the apparently reversed effect of context imageability. Namely, in some conditions, the test scores were higher for the less imageable context condition than the more imageable condition, although this kind of tendency was not statistically significant. This study presented a possible difference between the advanced learners and other learner groups in Section 6.2.2, although the amount of empirical data was quite limited. Future studies should examine whether advanced learners can take advantage of abstract contexts that do not describe a specific situation. If a condition where less imageable contexts enhance advanced learners' vocabulary recall is specified in future research, the cognitive strategies learners use while reading less imageable contexts should be also examined using the research paradigm proposed in Experiment 5.

Third, Section 6.2.3 points out the possibility that learners who use multiple knowledge sources at the same time are more sensitive to context imageability. Interestingly, in Experiment 4, the context-oriented learners were not as sensitive as the word-and-context-oriented ones, although this tendency was not statistically significant. However, as explained in Section 6.2.3, there were at least three different accounts for this tendency (i.e., explanation in terms of learner's cognitive resource, mental process connecting the word- and context-based information, and interference of other learner variables); this problem should be further investigated in future research. For example, a comparison of word-, context-, and word-and-context-oriented learner groups with similar English proficiency levels might be highly suggestive to test the plausibility of the third account. In

addition, an experiment comparing learning with and without a time limit will be also suggestive to differentiate the first and second accounts.

In addition to the three problems above, a general limitation of the present study was the small sample size in each experiment. A larger-scale study or meta-analysis of data might be quite informative. Furthermore, online research methods such as eye tracking will provide information that enables a more sophisticated discussion of the cognitive strategies used during learning a word list. For example, the time allocation during learning and learners' reading behavior such as rereading can be examined through eye tracking experiments. To examine the applicability of the current findings to other learning modes such as incidental vocabulary learning through reading and listening, it is necessary to study the effect of context reading focusing on different types of contexts and tasks. For example, it might be interesting to examine the context imageability effect when learners read multiple contexts instead of a single example sentence as in Ushiro et al. (2013).

7.3 Pedagogical Implications

This section explores the potential pedagogical implications of the findings in this research. This research clarified the advantages of contextualized vocabulary learning and the effective use of context. The two most important suggestions can be stated as follows. First, learners at a certain level of lexical proficiency may construct and retain the memory of context in relation to unfamiliar lexical items. If a teacher takes a contextualized approach to vocabulary instruction, it can be expected that students will learn the relationship between the target vocabulary and its context. They might thus learn how the meanings of words are conveyed and realized in context (Nation, 2001). Second, less imageable contexts are undesirable for contextualized vocabulary learning, except for advanced learners. The data from the current experiments implied that lower proficiency learners have difficulty utilizing the mental word–context association when the learning context is not very imageable.

Therefore, it is advisable for language teachers to consider whether students can imagine the situation described in the learning context when they are to take the contextualized approach to vocabulary instruction.

As reviewed in Section 2.5.2, researchers indicated that it is challenging for Japanese EFL learners to use their lexical knowledge flexibly. In particular, lower proficiency students seemed confused when they were asked to apply their knowledge to a new context. Considering that the conceptual representation is assumed to be gained through repeated exposures to contexts, the flexibility of lexical knowledge use might be cultivated through multiple encounters with context rather than a single contextualized task. In addition, it might be advisable for teachers to help their students at a lower level of proficiency to understand the contextual information and depict a mental image of the situation described in the context. This suggestion needs additional empirical support in terms of effective strategy use in vocabulary learning. Again, further research should be conducted taking into account learners' attention and cognitive strategies during learning.

With regard to multiple encounters with written context, Waring and Takaki (2003) revealed that Japanese learners gain greater vocabulary knowledge after reading a longer text. Their findings can be applied to the discussion of the effectiveness of extensive reading activity on lexical development. Past studies indicated that the effect of reading context on L2 vocabulary acquisition is enhanced by increasing exposure to context. One proposal of these studies was that vocabulary learning through reading is a gradual, incremental process where knowledge gains occur in small increments with repeated encounters. For example, Webb (2007b) showed how the productive and receptive knowledge of orthography, association, grammatical functions, syntax, and meaning/form was gained through 10 encounters with contextualized readings. The frequency of repetition in reading materials was considerably more influential on learning than the general frequency of words in the target language (Horst, Cobb, & Meara, 1998). In Horst et al. (1998), for example, participants showed no knowledge

gain of the meaning of *willing* and *harvest*, which were relatively frequent words in English. These words were used five times in a narrative story consisting of 21,232 words. In contrast, the participants successfully learned the words that were frequently used in the story such as *furmity*, which they encountered 12 times during reading. Based on this suggestion, the mental associations found in the present study might be fostered through repeated encounters.

Therefore, as assumed in Figure 7.1, a learner's mental lexicon should be enriched by multiple encounters in context because she or he can learn how the word meanings are realized in a variety of contexts. In terms of text genre, the material contexts in the present research were adapted from Webb (2008) and contained segments from easy narrative stories. Therefore, as in the extensive reading studies (Horst, 2005; Horst et al., 1998; Pigada & Schmitt, 2006), the rate of successful learning is expected to increase when students read an entire narrative story. Thus, extensive reading that provides students with multiple exposures to target words in meaningful contexts may be effective for development. In addition, in Oxford and Scarcella's (1994) words, fully contextualized activities such as extensive reading establish "links" between the instructed words and the students' own experiences and prior knowledge. For another example, Schouten-van Parreren (1989) reported that students often associate the instructed words with mental imagery when recalling word knowledge. In theory, learners' experiences and background knowledge enable their situation model representation through reading comprehension. The present study revealed that the imagery of a situation described in context establishes an association between the target words and related memories.

It should be noted again that the effective combination of intentional and incidental learning is important for L2 vocabulary learning. However, Folse's (2010) recent survey revealed that the amount of explicit vocabulary focus in an intensive English program was surprisingly low. In particular, a communication class focused on vocabulary eight times per week, while the vocabulary focus in a reading class was only three times a week. One such

vocabulary focus was the teacher's writing of 20 words from the week's story on the board and expecting the students to copy the vocabulary in preparation for a test the following week. Given the few opportunities for explicit instruction of vocabulary, it can be presumed that language teachers rarely consider that reading materials evoke mental imagery and this can support learner memory of new words. If it is difficult for teachers to focus on specific vocabulary in a reading class, an alternative way of encouraging the word-context association in learners' minds is to give instruction on deeper comprehension of the text. The mental representations created at the level of a situation model are an aspect of deeper comprehension that might considerably enhance incidental construction of the word-context association. However, as indicated in Experiment 2, instruction that only asks students to imagine the situation might not be enough. Instead, teachers should be more careful about whether the current context is comprehensible and imageable for the students.

In summary, the current research found that L2 learners can construct mental associations between unfamiliar words and learning contexts during reading under certain conditions. As mentioned in Section 2.5.2, Japanese EFL learners often utilize Japanese translation when they learn new words and read unfamiliar texts in English. For example, when using a dictionary or a word-list book to learn new words or to check the meaning of keywords used in a text, some students might read example sentences and others not. If learners' minds were not sensitive to contextual information at all, example sentences would not be necessary. Since past studies have not provided clear evidence showing the effect of context reading on the translation-based learning of new words, the present study explored the effect of context reading on intentional vocabulary learning, focusing on learners' imagery generation process. In constructing such imagery representations, less skilled students might need teachers' help because it is essential for them to comprehend the context fluently when they use contextual information for vocabulary learning (Griffin, 1992; Prince, 1996). This thesis hopefully demonstrates the importance of context use in learning L2 vocabulary.

References

Alba, J. W. (1984). Nature of inference representation. *American Journal of Psychology*, *97*, 215–233.

Alderson, J. C. (2000). *Assessing reading*. Cambridge, England: Cambridge University Press.

Anezaki, T. (2003). The effects of methods on EFL vocabulary learning: The role of contexts, collocations and translations. *Annual Review of English Language Education in Japan*, *14*, 171–180.

Antonietti, A. (1991). Why does mental visualization facilitate problem solving? In R. H. Logie & M. Denis (Eds.), *Mental images in human cognition* (pp. 211–228). Amsterdam, the Netherlands: Elsevier.

Antonietti, A. (1999). Can students predict when imagery will allow them to discover the problem solution? *European Journal of Cognitive Psychology*, *11*, 407–428.

Baleghizadeh, S., & Shahry, M. N. N. (2011). The effect of three consecutive context sentences on EFL vocabulary-learning. *TESL Canada Journal*, *28*, 74–89.

Barcroft, J. (2002). Semantic and structural elaboration in L2 lexical acquisition. *Language Learning*, *52*, 323–363.

Barcroft, J. (2003). Effects of questions about word meaning during L2 Spanish lexical learning. *The Modern Language Journal*, *87*, 546–561.

Barcroft, J. (2004). Effects of sentence writing in second language lexical acquisition. *Second Language Research*, *20*, 303–334.

Begg, I., & Paivio, A. (1969). Concreteness and imagery in sentence meaning. *Journal of Verbal Learning and Verbal Behavior*, *8*, 821–827.

Block, E. (1986). The comprehension strategies of second language readers. *TESOL Quarterly*, *20*, 463–494.

Bolger, D. J., Balass, M., Landen, E., & Perfetti, C. A. (2008). Context variation and

definitions in learning the meanings of words: An instance-based learning approach. *Discourse Processes*, *45*, 122–159.

Bruton, A., Lopez, M. G., & Mesa, R. E. (2011). Incidental L2 vocabulary learning: An impracticable term? *TESOL Quarterly*, *45*, 759–768.

Carrell, P. L. (1991). Awareness of text structure: Effects on recall. *Language Learning*, *41*, 1–20.

Carver, R. P. (1992). Reading rate: Theory, research, and practical implications. *Journal of Reading*, *36*, 84–95.

Chaffin, R., Morris, R. K., & Seely, R. E. (2001). Learning new word meanings from context: A study of eye movements. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *27*, 225–235.

Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology*, *33*, 497–505.

Committee of revising the JACET basic words (Ed.). (2003). *JACET list of 8000 basic words*. Tokyo, Japan: the Japan Association of College English Teachers.

Cook, G. (2010). *Translation in language teaching*. Oxford, England: Oxford University Press.

Cook, V. (1997). The consequences of bilingualism for cognitive processing. In A. M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 279–299). Mahwah, NJ: Erlbaum.

Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge, England: Cambridge University Press.

Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions*. New York, NY: Irvington.

Date, M. (2006). The influence of ways to trigger and enhance awareness in incidental learning of new vocabulary. *Annual Review of English Language Education in Japan*, *17*,

131–140.

Davey, B. (1988). Factors affecting the difficulty of reading comprehension items for successful and unsuccessful readers. *Journal of Experimental Education*, *56*, 67–76.

Day, R. R., Omura, C., & Hiramatsu, M. (1991). Incidental EFL vocabulary learning and reading. *Reading in a Foreign Language*, *7*, 541–551.

de Bot, K., Paribakht, T. S., & Wesche, M. B. (1997). Toward a lexical processing model for the study of second language vocabulary acquisition. *Studies in Second Language Acquisition*, *19*, 309–329.

de Groot, A. M. B. (1992). Bilingual lexical representation: A closer look at conceptual representations. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 389–412). Amsterdam, the Netherlands: Elsevier.

de Groot, A. M. B. (2011). *Language and cognition in bilinguals and multilinguals: An introduction.* New York, NY: Psychology Press.

de Groot, A. M. B., & Keijzer, R. (2000). What is hard to learn is easy to forget: The roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Language Learning*, *50*, 1–56.

de Groot, A. M. B., & Poot, R. (1997). Word translation at three levels of proficiency in a second language: The ubiquitous involvement of conceptual memory. *Language Learning*, *47*, 215–264.

Denis, M. (1989). *Les images mentales* [The mental imagery]. (R. Terauchi, Trans.). Tokyo, Japan: Keiso Shobo. (Original work published 1979).

Dollerup, C., Glahn, E., & Hansen, C. R. (1989). Vocabularies in the reading process. In P. Nation & R. Carter (Eds.), *Vocabulary acquisition* (pp. 21–33). Amsterdam, the Netherlands: Free University Press.

Dunlea, J. (2009). *Eiken to CEFR no kanren-sei ni tsuite* [About the relationship between Eiken and the CEFR]. Retrieved from https://www.eiken.or.jp/ eiken/group/result/

Dupuy, B., & Krashen, S. D. (1993). Incidental vocabulary acquisition in French as a foreign language. *Applied Language Learning*, *4*, 55–63.

Duthie, J. K., Nippold, M. A., Billow, J. L., & Mansfield, T. C. (2008). Mental imagery of concrete proverbs: A developmental study of children, adolescents, and adults. *Applied Psycholinguistics*, *29*, 151–173.

Educational Testing Service (2006). *The official guide to the new TOEFL iBT*. New York, NY: McGraw-Hill.

Educational Testing Service. (2010). *Eiken shutoku-sha no TOEIC® score* [TOEIC® scores among STEP Eiken successful examinees]. Retrieved from http://www.toeic.or.jp/toeic/pdf/data/TOEIC-STEP_2001.pdf

Elgort, I. (2011). Deliberate learning and vocabulary acquisition in a second language. *Language Learning*, *61*, 367–413.

Elgort, I., & Warren, P. (2014). L2 vocabulary learning from reading: Explicit and tacit lexical knowledge and the role of learner and item variables. *Language Learning*, *64*, 365–414.

Ellis, N. C., & Beaton, A. (1993a). Factors affecting the learning of foreign language vocabulary: Imagery keyword mediators and phonological short-term memory. *Quarterly Journal of Experimental Psychology*, *46*, 533–558.

Ellis, N. C., & Beaton, A. (1993b). Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning*, *43*, 559–617.

Ellis, R. (2008). *The study of second language acquisition* (2nd ed.). Oxford, England: Oxford University Press.

Ellis, R., & Shintani, N. (2014). *Exploring language pedagogy through second language acquisition research*. London, England: Routledge.

Fan, M. Y. (2003). Frequency of use, perceived usefulness, and actual usefulness of second language vocabulary strategies: A study of Hong Kong learners. *The Modern Language Journal*, *87*, 222–241.

Fincher-Kiefer, R., & D'Agostino, P. R. (2004). The role of visuospatial resources in generating predictive and bridging inferences. *Discourse Processes*, *37*, 205–224.

Folse, K. (2004). *Vocabulary myths: Applying second language research to classroom teaching.* Ann Arbor, MI: University of Michigan Press.

Folse, K. (2010). Is explicit vocabulary focus the reading teachers' job? *Reading in a Foreign Language*, *22*, 139–160.

Franco-Watkins, A. M., & Dougherty, M. (2006). Word concreteness and encoding effects on context-dependent discrimination. *Memory & Cognition*, *34*, 973–985.

Fukkink, R. G., Blok, H., & de Glopper, K. (2001). Deriving word meaning from context: A multicomponential skill. *Language Learning*, *51*, 477–496.

Gambrell, L. B., & Jawitz, P. B. (1993). Mental imagery, text illustrations, and children's story comprehension and recall. *Reading Research Quarterly*, *28*, 265–276.

Godfroid, A., Boers, F., & Housen, A. (2013). An eye for words: Gauging the role of attention in incidental L2 vocabulary acquisition by means of eye-tracking. *Studies in Second Language Acquisition*, *35*, 483–517.

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

Griffin, G. F. (1992). *Aspects of the psychology of second language vocabulary list learning.* Doctoral dissertation, University of Warwick, England. Retrieved from http://wrap. warwick.ac.uk/36070/

Griffin, Z. M., & Bock, K. (1998). Constraint, word frequency, and the relationship between lexical processing levels in spoken word production. *Journal of Memory and Language*, *38*, 313–338.

Hasegawa, Y. (2010). Context effects on EFL vocabulary learning: Focusing on word imageability and test formats. *JLTA (the Japan Language Testing Association) Journal*, *13*, 145–161.

Hasegawa, Y. (2013). Who will benefit from glossed context in intentional vocabulary learning? Comparing high school and university students. *JLTA Journal*, *16*, 43–64.

Horiba, Y. (1996). Comprehension processes in L2 reading: Language competence, textual coherence, and inferences. *Studies in Second Language Acquisition*, *18*, 433–473.

Horiba, Y. (2002). The effect of task on comprehension and memory of L1 and L2 text. *Scientific Approaches to Language*, *1*, 35–47.

Horiba, Y. (2013). Task-induced strategic processing in L2 text comprehension. *Reading in a Foreign Language*, *25*, 98–125.

Horst, M. (2005). Learning L2 vocabulary through extensive reading: A measurement study. *The Canadian Modern Language Review*, *61*, 355–382.

Horst, M., Cobb, T., & Meara, P. (1998). Beyond a Clockwork Orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language*, *11*, 207–223.

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

Hulstijn, J. H. (1992). Retention of inferred and given word meaning: Experiments in incidental vocabulary learning. In P. Arnaud & H. Béjoint (Eds.), *Vocabulary and applied linguistics* (pp. 102–112). London, England: Macmillan.

Hulstijn, J. H. (2005). Theoretical and empirical issues in the study of implicit and explicit second-language learning. *Studies in Second Language Acquisition*, *27*, 129–140.

Hulstijn, J. H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning*, *51*, 539–558.

Hunt, A., & Beglar, D. (2005). A framework for developing EFL reading vocabulary. *Reading in a Foreign Language*, *17*, 23–59.

Ikeno, O. (1996). The effects of text-structure-guiding questions on comprehension of texts with varying linguistic difficulties. *JACET (Japan Association of College English Teachers) Bulletin*, *27*, 51–68.

Jiang, N. (2000). Lexical representation and development in a second language. *Applied Linguistics*, *21*, 47–77.

Joe, A. (2010). The quality and frequency of encounters with vocabulary in an English for Academic Purposes programme. *Reading in a Foreign Language*, *22*, 117–138.

Just, M. A., Newman, S. D., Keller, T. A., McEleney, A., & Carpenter, P. A. (2004). Imagery in sentence comprehension: An fMRI study. *NeuroImage*, *21*, 112–124.

Kaakinen, J. K., & Hyönä, J. (2005). Perspective effects on expository text comprehension: Evidence from think-aloud protocols, eyetracking, and recall. *Discourse Processes*, *40*, 239–257.

Kadota, S. (2006) *Daini-gengo rikai no ninchi mechanism: Eigo no kaki-kotoba no shori to on'in no yakuwari* [Reading and phonological processes in English as a second language]. Tokyo, Japan: Kurosio Publishers.

Kadota, S., Noro, T., Shiki, O., & Hase, N. (Eds.). (2014). *Eitango un'you-ryoku hantei soft o tsukatta goi-shidou* [English vocabulary instruction using lexical proficiency assessment software]. Tokyo, Japan: Taishukan.

Kimura, Y. (1999). Metacognitive awareness training and reading comprehension of Japanese EFL learners. *JACET Bulletin*, *30*, 45–58.

Kojic-Sabo, I., & Lightbown, P. M. (1999). Students' approaches to vocabulary learning and their relationship to success. *The Modern Language Journal*, *83*, 176–192.

Konishi, T., & Minamide, K. (Eds.). (2001). *Taishukan's genius unabridged English-Japanese dictionary*. Tokyo, Japan: Taishukan.

Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *The Modern Language Journal*, *73*, 440–464.

Krasny, K. A., & Sadoski, M. (2008). Mental imagery and affect in English/French bilingual readers: A cross-linguistic perspective. *The Canadian Modern Language Review*, *64*, 399–428.

Kroll, J. F., & de Groot, A. M. B. (1997). Lexical and conceptual memory in the bilingual: Mapping form to meaning in two languages. In A. M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 169–199). Mahwah, NJ: Erlbaum.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, *33*, 149–174.

Kroll, J. F., van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The revised hierarchical model: A critical review and assessment. *Bilingualism: Language and Cognition*, *13*, 373–381.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Larson-Hall, J. (2010). *A guide to doing statistics in second language research using SPSS*. New York, NY: Routledge.

Laufer, B., Elder, C., Hill, K., & Congdon, P. (2004). Size and strength: Do we need both to measure vocabulary knowledge? *Language Testing*, *21*, 202–226.

Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language Learning*, *54*, 399–436.

Laufer, B., & Hulstijn, J. H. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, *22*, 1–26.

Laufer, B., & Nation, I. S. P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, *16*, 307–322.

Laufer, B., & Nation, I. S. P. (2012). Vocabulary. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 163–176). New York, NY: Routledge.

Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, *22*, 15–30.

Laufer, B., & Shmueli, K. (1997). Memorizing new words: Does teaching have anything to do with it? *RELC Journal*, *28*, 89–108.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). Introduction to Latent Semantic Analysis. *Discourse Processes*, *25*, 259-284.

Lawson, M. J., & Hogben, D. (1996). The vocabulary-learning strategies of foreign-language students. *Language Learning*, *46*, 101–135.

Maeda, H., Tagashira, K., & Miura, H. (2003). Vocabulary learning strategy use and learning achievement by Japanese high school EFL learners. *Japanese Journal of Educational Psychology*, *51*, 273–280.

Ministry of Education, Culture, Sports, Science and Technology. (2008). *Chugakkou gakushu shidou youryou* [The Course of Study for lower secondary school]. Retrieved from http://www.mext.go.jp/a_menu/shotou/new-cs/ youryou/chu/index.htm

Ministry of Education, Culture, Sports, Science and Technology. (2009). *Koutou-gakkou gakushu shidou youryou* [The Course of Study for upper secondary school]. Retrieved from http://www.mext.go.jp/a_menu/shotou/ new-cs/youryou/ 1304427.htm

Mizumoto, A., & Takeuchi, O. (2009). Examining the effectiveness of explicit instruction of vocabulary learning strategies with Japanese EFL university students. *Language Teaching Research*, *13*, 425–449.

Mochizuki, M., Aizawa, K., & Tono, Y. (2003). *Eigo-goi no shidou manual* [The teaching manual of English vocabulary]. Tokyo, Japan: Taishukan.

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.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, *16*, 519–533.

Murnane, K., Phelps, M. P., & Malmberg, K. (1999). Context-dependent recognition memory: The ICE theory. *Journal of Experimental Psychology: General*, *128*, 403–415.

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

Nassaji, H. (2003). L2 vocabulary learning from context: Strategies, knowledge sources, and their relationship with success in L2 lexical inferencing. *TESOL Quarterly*, *37*, 645–670.

Nassaji, H. (2006). The relationship between depth of vocabulary knowledge and L2 learners' lexical inferencing strategy and use. *The Modern Language Journal*, *90*, 387–401.

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

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

Nation, I. S. P. (2005). Teaching and learning vocabulary. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 581–595). Mahwah, NJ: Erlbaum.

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

Nation, I. S. P. (2009). *Teaching vocabulary: Strategies and techniques*. Cambridge, England: Cambridge University Press.

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

Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *Language Teacher*, *31*, 9–13.

Nation, I. S. P., & Webb, S. (2011). *Researching and analyzing vocabulary*. Boston, MA: Heinle.

Ogawa, T., & Inamura, Y. (1974). An analysis of word attributes imagery, concreteness, meaningfulness and ease of learning for Japanese nouns. *The Japanese Journal of Psychology*, *44*, 317–327.

O'Loughlin, K. J. (1995). Lexical density in candidate output on direct and semi-direct versions of an oral proficiency test. *Language Testing*, *12*, 217–237.

O'Loughlin, K. J. (1997). *The comparability of direct and semi-direct speaking tests: A case study*. Unpublished Ph.D. thesis, University of Melbourne.

Oxford, R. L., & Scarcella, R. C. (1994). Second language vocabulary learning among adults: State of the art in vocabulary instruction. *System*, *22*, 231–243.

Paivio, A. (1986). *Mental representations: A dual coding approach.* Oxford, England: Oxford University Press.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, *45*, 255–287.

Paivio, A. (2006). *Mind and its evolution: A dual coding theoretical approach*. Mahwah, NJ: Erlbaum.

Peters, E., Hulstijn, J. H., & Sercu, L. (2009). Learning L2 German vocabulary through reading: The effect of three enhancement techniques compared. *Language Learning*, *59*, 113–151.

Pickering, M. (1982). Context-free and context-independent vocabulary learning: An experiment. *System*, *10*, 79–83.

Pigada, M., & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language*, *18*, 1–27.

Pitts, M., 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*, 271–275.

Prado, E. L., & Ullman, M. T. (2009). Can imageability help us draw the line between storage and composition? *Journal of Experimental Psychology: Learning, Memory and Cognition*, *35*, 849–866.

Prince, P. (1996). Second language vocabulary learning: The role of context versus translations as a function of proficiency. *The Modern Language Journal*, *80*, 478–493.

Qian, D. D. (2008). From single words to passages: Contextual effects on predictive power of vocabulary measures for assessing reading performance. *Language Assessment Quarterly*, *5*, 1–19.

Read, J. (2000). *Assessing vocabulary.* Cambridge, England: Cambridge University Press.

Read, J., & Chapelle, C. A. (2001). A framework for second language vocabulary assessment. *Language Testing*, *18*, 1–32.

Richardson, J. T. E. (2002). *Imagery*. (T. Nishimoto, Trans.). Tokyo, Japan: Waseda University Press. (Original work published 1999).

Rott, S. (2005). Processing glosses: A qualitative exploration of how form-meaning connections are established and strengthened. *Reading in a Foreign Language*, *17*, 95–124.

Rowe, E. J., Schurr, B., Meisinger, D. (1978). Comprehensibility ratings of concrete and abstract sentences. *Bulletin of the Psychonomic Society*, *11*, 49–52.

Sadoski, M., Goetz, E. T., & Fritz, J. B. (1993). A causal model of sentence recall: Effects of familiarity, concreteness, comprehensibility, and interestingness. *Journal of Reading Behavior*, *25*, 5–16.

Sadoski, M., Goetz, E. T., & Rodriguez, M. (2000). Engaging texts: Effects of concreteness on comprehensibility, interest, and recall in four text types. *Journal of Educational Psychology*, *92*, 85–95.

Sadoski, M. & Paivio, A. (2004). A dual coding theoretical model of reading. In R. B. Ruddell & N. J. Unrau (Eds.), *Theoretical models and processes of reading* (5th ed., pp. 1329–

1362). Newark, DE: International Reading Association.

Sakuma, N., Ijuin, M., Fushimi, T., Tatsumi, I., Tanaka, M., Amano, S., & Kondo, K. (2005). *NTT database series. Nihongo no goi-tokusei: Tango-shinzousei database* [Japanese vocabulary attribution: Word imageability database]. Tokyo, Japan: Sanseido.

Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research*, *12*, 329–363.

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual.* New York, NY: Palgrave Macmillan.

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

Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching*, *47*, 484–503.

Schouten-van Parreren, C. (1989). Vocabulary learning through reading: Which conditions should be met when presenting words in texts? In P. Nation & R. Carter (Eds.), *Vocabulary acquisition* (pp. 75–85). Amsterdam, the Netherlands: Free University Press.

Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *Journal of Memory and Language*, *55*, 197–212.

Shirahata, T., Tomita, Y., Muranoi, H., & Wakabayashi, S. (2009). *Eigo-kyoiku yougo jiten: Kaitei-ban* [A guide to English language teaching terminology: Revised edition]. Tokyo, Japan: Taishukan.

Society for Testing English Proficiency [STEP]. (2007). *The EIKEN can-do list: English translation*. Retrieved from https://www.eiken.or.jp/eiken/exam/ cando/

Society for Testing English Proficiency [STEP]. (2010). *Eiken test in practical English proficiency*. Retrieved from http://www.eiken.or.jp/

Sonbul, S., & Schmitt, N. (2010). Direct teaching of vocabulary after reading: Is it worth the effort? *ELT Journal*, *64*, 253–260.

Sonbul, S., & Schmitt, N. (2013). Explicit and implicit lexical knowledge: Acquisition of collocations under different input conditions. *Language Learning*, *63*, 121–159.

Squires, D. A., Huitt, W. G., and Segars, J. K. (1983). *Effective schools and classrooms: A research-based perspective*. Alexandria. VA: Association for Supervisor and Curricular Development.

Steinel, M. P., Hulstijn, J. H., & Steinel, W. (2007). Second language idiom learning in a paired-associate paradigm: Effects of direction of learning, direction of testing, idiom imageability, and idiom transparency. *Studies in Second Language Acquisition*, *29*, 449–484.

Stenner, A. J. (1996). *Measuring reading comprehension with the Lexile Framework*. Paper presented at the 4th North American Conference on Adolescent/Adult Literacy, Washington, D.C.

Stenner, A. J., Burdick, H., Sanford, E. E., & Burdick, D. S. (2006). How accurate are Lexile text measures? *Journal of Applied Measurement*, *7*, 307–322.

Sunderman, G., & Kroll, J. F. (2006). First language activation during second language lexical processing: An investigation of lexical form, meaning, and grammatical class. *Studies in Second Language Acquisition*, *28*, 387–422.

Swanborn, M. S. L., & de Glopper, K. (1999). Incidental word learning while reading: A meta-analysis. *Review of Educational Research*, *69*, 261–285.

Swanborn, M. S. L., & de Glopper, K. (2002). Impact of reading purpose on incidental word learning from context. *Language Learning*, *52*, 95–117.

Tapiero, I. (2007). *Situation models and levels of coherence.* Mahwah, NJ: Erlbaum.

Thorndyke, P. W. (1977). Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, *9*, 77–110.

Tono, Y. (Ed.). (2012). *The final report of the CEFR-J project*. A report of a Grant-in-Aid for Scientific Research No. 20242011. Retrieved from http://cefr-j.org/ research.html

Tono, Y. (2013). *Eigo toutatsu-do shihyou CEFR-J guidebook: Can-do list sakusei/katsuyou* [The CEFR-J handbook: A resource book for using can-do descriptors for English language teaching]. Tokyo, Japan: Taishukan.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*, 352–373.

Ushiro, Y., Hasegawa, Y., Nahatame, S., Hamada, A., Kimura, Y., Mori, Y., & Kato, D. (2013). Incremental learning of homonyms in multiple contexts among Japanese EFL readers. *JACET Journal*, *57*, 1–20.

Ushiro, Y., Hasegawa, Y., Nahatame, S., Shimizu, H., Takaki, S., Hamada, A., & Nakagawa, C. (2012). How Japanese EFL readers revise their situation models: Focusing on reading skills and the causal structure of texts. *Annual Review of English Language Education in Japan*, *23*, 105–120.

Ushiro, Y., Hoshino, Y., Shimizu, H., Kai, A., Nakagawa, C., Watanabe, F., & Takaki, S. (2010). Disambiguation of homonyms by EFL readers: The effects of primary meaning and context interpretation. *Annual Review of English Language Education in Japan*, *21*, 161–170.

Ushiro, Y., Kai, A., Shimizu, H., Hoshino, Y., Nahatame, S., Hasegawa, Y., Yano, K., & Nakagawa, C. (2011). Effects of flashback on Japanese EFL readers' narrative comprehension. *Annual Review of English Language Education in Japan*, *22*, 111–126.

van Assche, E., Drieghe, D., Duyck, W., Welvaert, M., & Hartsuiker, R. J. (2011). The influence of semantic constraints on bilingual word recognition during sentence reading. *Journal of Memory and Language*, *64*, 88–107.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension.* New York, NY: Academic Press.

van Hell, J. G., & de Groot, A. M. B. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition*, *1*, 193–211.

van Hell, J. G., & de Groot, A. M. B. (2008). Sentence context modulates visual word recognition and translation in bilinguals. *Acta Psychologica*, *128*, 431–451.

Vatz, K., Tare, M., Jackson, S. R., & Doughty, C. J. (2013). Aptitude-treatment interaction studies in second language acquisition: Findings and methodology. In G. Granena & M. H. Long (Eds.), *Sensitive periods, language aptitude, and ultimate L2 attainment* (pp. 273–292). Amsterdam, the Netherlands: John Benjamins.

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*, 130–163.

Webb, S. (2007a). Learning word pairs and glossed sentences: The effects of a single context on vocabulary knowledge. *Language Teaching Research*, *11*, 63–81.

Webb, S. (2007b). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, *28*, 46–65.

Webb, S. (2008). The effects of context on incidental vocabulary learning. *Reading in a Foreign Language*, *20*, 232–245.

Wesche, M., & Paribakht, T. S. (1996). Assessing second language vocabulary knowledge: Depth versus breadth. *The Canadian Modern Language Review*, *53*, 13–40.

Wesche, M., & Paribakht, T. S. (2010). *Lexical inferencing in a first and second language: Cross-linguistic dimensions*. Bristol, England: Multilingual Matters.

Woollams, A. M. (2005). Imageability and ambiguity effects in speeded naming: Convergence and divergence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 878–890.

Yamashita, J. (2011). Hatsuwa-protocol o tsukatta ninchi-process no kenkyu [Studies of cognitive processes with think-aloud protocols]. In I. Fujimura & N. Takizawa (Eds.),

Gengo-kenkyu no giho: Data no shushu to bunseki [Techniques for linguistic studies: Data collection and analysis] (pp. 111–121). Tokyo, Japan: Hituzi Syobo.

Yokokawa, H. (Ed.). (2006). *Nihonjin eigo gakushusha no eitango shinmitsudo (Moji hen): Kyouiku kenkyu no tame no dai-ni gengo database* [English word familiarity of Japanese EFL learners: L2 database for research and education]. Tokyo, Japan: Kuroshio.

Yoshida, M. (2003). Working memory capacity and the use of inference in L2 reading. *JACET Bulletin*, *36*, 1–17.

Zahar, R., Cobb, T., & Spada, N. (2001). Acquiring vocabulary through reading: Effects of frequency and contextual richness. *The Canadian Modern Language Review*, *57*, 541–572.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*, 162–185.

Appendices

Appendix 3.1

The Target Words, Translations, and Contexts Used in Pilot Study 1

Target (meaning)	Translation		Context	INF	IMG
ancon (hospital)	病院 “byoin”	1	He was not ill, and of course the beds in the ancon are for ill people.	4	4.11
		2	As soon as he could walk, he left the ancon and started looking for a ship to take him back to England.	2	6.00
		3	I did not talk to him very much at the ancon. I looked at his head and arms and legs and body very carefully.	2	5.89
		4	I am a doctor at the London Ancon.	4	6.56
		5	Time passed slowly in the ancon, where the patients played cards, and slept, and told each other stories.	4	6.00
cader (lunch)	昼食 “chushoku”	1	I’ve ordered some champagne and some cader for us.	3	5.56
		2	On her last day, Mary and John had cader in the factory restaurant together.	3	6.33
		3	Paul found the Minister sitting alone in the garden after cader, smoking a large cigar, his big red hands folded over his large stomach, a soft hat over his eyes.	2	4.78
		4	He had cader and tea with Mrs. Walsh in the kitchen, and he felt really happy.	3	6.22
		5	After cader we sat and talked for a while.	2	6.20
dangy (street)	街の通り “machi no tori”	1	Children played in the narrow dangies.	2	6.22
		2	So John had to walk back through the dangies in the middle of the night, all wet from the sea, while we laughed at him from the car.	2	3.56
		3	After paying the driver and getting out, he suddenly realized how dark and dirty the narrow dangies were.	3	5.56
		4	At the corner of the dangy she met Laurie.	2	6.44
		5	Big white clouds over the white dangies—and sunshine everywhere.	1	4.70

Appendix 3.1 (*continued*)

Target (meaning)	Translation		Context	INF	IMG
denent (remember)	～を覚えている “~ o oboete-iru”	1	Perhaps this happened. Or perhaps she’s ill and can’t denent who she is.	2	3.56
		2	We liked to talk to him, and we are all very sorry because he is dead. A lot of people are going to denent him for a long time.	4	5.33
		3	Don’t you denent you promised you would never leave me?	4	4.89
		4	Then Simon stood up to say something. “Mr Duncan,” he said. “I’ll always denent this night.”	3	3.89
		5	Perhaps I had a Christmas with my mother once, but I do not denent it.	3	6.30
faddam (write)	～を書く “~ o kaku”	1	She asked a lot of people a lot of questions and always faddamed the answers in her notebook.	4	5.78
		2	Your brother faddamed his name on this paper.	3	6.89
		3	I want to read what she faddamed to her dear husband.	3	5.78
		4	“Why don’t you faddam another book?” said Archie.	3	5.56
		5	He can read and faddam, and he thinks a lot.	4	5.80
hodet (face)	顔 “kao”	1	His hodet was shining, excited, happy.	2	6.22
		2	His hodet was red and he looked at his shoes.	2	5.44
		3	I could not see his hodet or his body.	3	5.89
		4	The first time Mary and the children laughed together, John felt a big smile come onto his hodet.	4	5.44
		5	As she stood up to see if her bag was all right, she saw her hodet in the mirror—white, with big, round eyes.	4	4.60
masco (train)	列車 “ressha”	1	Agatha still only eighteen years old went to London on the masco.	3	5.67
		2	And in London, hundreds of people were waiting at King’s Cross Station for the masco from Harrogate.	4	6.00
		3	Everyone on the masco and the ship looked at him, and laughed at him.	3	6.00
		4	The masco left the station and rushed into the dark.	2	6.22

Appendix 3.1 (*continued*)

Target (meaning)	Translation		Context	INF	IMG
		5	The masco began to go more slowly. It gave a long whistle. They were coming to a town.	3	5.40
pacon (wear)	～を身につけている “~ o mini-tsukete-iru”	1	They were all paconing white gloves and their hands were inside a big glass box.	3	5.00
		2	He insisted on paconing a bright yellow sports jacket and diamond tie-pin while serving lunch.	4	4.44
		3	Their friends were paconing white clothes, with Greenworld written on them.	3	4.56
		4	I don’t know anything about art, and I haven’t met any grand people, and I don’t pacon expensive clothes, but up to now, none of that has worried me.	3	3.78
		5	She was paconing a white coat with a lot of pens in the top pocket.	3	6.10
sagod (visit)	～を訪問する “~ o homon-suru”	1	Her brother Edward was always with her when the King came to sagod.	1	5.44
		2	I don’t want to be rich, but I do want to come to Australia and sagod you and Mollie and my grandchildren.	4	5.89
		3	Then a cat came to sagod me—a beautiful cat—and then a dog.	1	5.11
		4	I sagoded him every day, and talked to him.	3	6.22
		5	She remembered seeing a beautiful young girl in a hotel in Cairo when she was sagoding Egypt with Clara.	3	5.80
tasper (evening)	夕方 “yugata”	1	Bob Tappin and Bob Leeming played music in the hotel each tasper, and both of them watched the quiet woman in the corner of the room and began to think.	3	4.56
		2	During the tasper, she talked to some people who were just back from Baghdad, in Iraq.	2	5.33
		3	There are no long taspers in our New Zealand days; the sun goes down and half an hour later it’s night.	4	4.89
		4	They danced together many times that tasper, and Archie told her his plans.	2	3.89
		5	Archie stayed for the rest of the afternoon, and for supper that tasper.	4	4.30

Note. INF = the informativeness ratings in Webb (2008). IMG = the mean imageability ratings.

Appendix 3.2

The Format of the Imageability Rating Task Used in Pilot Study 1

英文をよく読んで、それぞれの文がどれくらいイメージしやすいかを 7 段階で評価してください。英文を読んで鮮明なイメージをすばやく思い浮かべることができる場合は「7」、まったくイメージを思い浮かべることができない場合は「1」に○をつけてください。

【例】Mrs. Smith was a very pious woman who attended church services regularly.
※pious =「信心深い」
（イメージしづらい） 1 2 3 4 5 ⑥ 7 （イメージしやすい）

なお、文中の難しい単語には和訳をつけてあります。

Her brother Edward was always with her when the King came to sagod.
※sagod =「～を訪問する」
（イメージしづらい） 1 2 3 4 5 6 7 （イメージしやすい）

He was not ill, and of course the beds in the ancon are for ill people.
※ancon =「病院」
（イメージしづらい） 1 2 3 4 5 6 7 （イメージしやすい）

His hodet was shining, excited, happy.
※hodet =「顔」
（イメージしづらい） 1 2 3 4 5 6 7 （イメージしやすい）

Children played in the narrow dangies.
※dangy =「（町の）通り」
（イメージしづらい） 1 2 3 4 5 6 7 （イメージしやすい）

I've ordered some champagne and some cader for us.
※cader =「昼食」
（イメージしづらい） 1 2 3 4 5 6 7 （イメージしやすい）

Bob Tappin and Bob Leeming played music in the hotel each tasper, and both of them watched the quiet woman in the corner of the room and began to think.
※tasper =「夕方」
（イメージしづらい） 1 2 3 4 5 6 7 （イメージしやすい）

She asked a lot of people a lot of questions and always faddamed the answers in her notebook.
※faddam =「～を書く」
（イメージしづらい） 1 2 3 4 5 6 7 （イメージしやすい）

Perhaps this happened. Or perhaps she's ill and can't denent who she is.
※denent =「～を覚えている」
（イメージしづらい） 1 2 3 4 5 6 7 （イメージしやすい）

They were all paconing white gloves and their hands were inside a big glass box.
※pacon =「～を身につけている」
（イメージしづらい） 1 2 3 4 5 6 7 （イメージしやすい）

Agatha still only eighteen years old went to London on the masco.
※masco =「列車」
（イメージしづらい） 1 2 3 4 5 6 7 （イメージしやすい）

指示があるまで次のページに進まないでください。

Appendix 3.3

The Format of the Context Recall Task Used in Pilot Study 1

先ほど読んだ英文に出てきた以下の単語をみて、それぞれの単語がどのような文の中で使われていたかを思い出して、できるだけ多くの内容を日本語で書き出してください。

【例】 pious (信心深い)

スミス夫人は、いつも礼拝に参加する、とても信心深い女性だった。

(1) ancon (病院)

(2) cader (昼食)

(3) dangy (（町の）通り)

(4) denent (～を覚えている)

(5) faddam (～を書く)

(6) hodet (顔)

(7) masco (列車)

(8) pacon (～を身につけている)

(9) sagod (～を訪問する)

(10) tasper (夕方)

指示があるまで次のページに進まないでください。

Appendix 3.4

Examples of the Target Words, Translations, and Contexts Used in Experiment 1

Target (meaning)	Translation	Context	Imageability
ancon (hospital)	病院 “byoin”	I am a doctor at the London Ancon.	more
cader (lunch)	昼食 “chushoku”	Paul found the Minister sitting alone in the garden after cader, smoking a large cigar, his big red hands folded over his large stomach, a soft hat over his eyes.	less
dangy (street)	街の通り “machi no tori”	At the corner of the dangy she met Laurie.	more
denent (remember)	～を覚えている “~ o oboete-iru”	Perhaps this happened. Or perhaps she’s ill and can’t denent who she is.	less
faddam (write)	～を書く “~ o kaku”	Your brother faddamed his name on this paper.	more
hodet (face)	顔 “kao”	As she stood up to see if her bag was all right, she saw her hodet in the mirror—white, with big, round eyes.	less
masco (train)	列車 “ressha”	The masco left the station and rushed into the dark.	more
pacon (wear)	～を身につけている “~ o mini-tsukete-iru”	I don’t know anything about art, and I haven’t met any grand people, and I don’t pacon expensive clothes, but up to now, none of that has worried me.	less
sagod (visit)	～を訪問する “~ o homon-suru”	I sagoded him every day, and talked to him.	more
tasper (evening)	夕方 “yugata”	They danced together many times that tasper, and Archie told her his plans.	less

Appendix 3.5

The Format of the Contextualized Recognition Test Used in Experiment 1

4 つの選択肢の中から、英文中の空欄に最もよくあてはまるものを 1 つ選んで、丸をつけて解答してください。解答時間は最大で 5 分間（1 問あたり 30 秒）です。

■Q1

Children played in the narrow (　　).

①caders　②copacs　③dangies　④tagons

■Q2

They were all (　　) white gloves and their hands were inside a big glass box.

①paconing　②nasining　③taspering　④mesutting

■Q3

Time passed slowly in the (　　), where the patients played cards, and slept, and told each other stories.

①faddam　②copac　③nasin　④ancon

■Q4

I could not see his (　　) or his body.

①ancon　②gishom　③hodet　④hattaw

■Q5

We liked to talk to him, and we are all very sorry because he is dead. A lot of people are going to (　　) him for a long time.

①nuggy　②denent　③toncop　④dangy

■Q6

Everyone on the (　　) and the ship looked at him, and laughed at him.

①sagod　②nuggy　③masco　④ictay

■Q7

There are no long (　　) in our New Zealand days; the sun goes down and half an hour later it's night.

①tagons　②taspers　③ictays　④hodets

■Q8

He can read and (　　), and he thinks a lot.

①faddam　②dapew　③gishom　④masco

■Q9

He had (　　) and tea with Mrs. Walsh in the kitchen, and he felt really happy.

①toncop　②mesut　③cader　④denent

■Q10

I don't want to be rich, but I do want to come to Australia and (　　) you and Mollie and my grandchildren.

①dapew　②hattaw　③pacon　④sagod

指示があるまで次のページに進まないでください。

Appendix 3.6

Examples of the Target Words, Translations, and Contexts Used in Experiments 2, 4 and 6

Target	Translation	Context	Imageability
apprentice	見習い “minarai”	Before John became a professional chef, he spent many years working as an apprentice to a master chef at a famous restaurant in London.	more
assorted	各種取りそろえた “kakushu tori-soroeta”	A: Darling, what kind of cheese should we serve at our dinner party on Saturday? B: I’m going to prepare a tray with assorted cheeses on it. That way, people can choose what they like.	more
clutch	ぐいと握る “gui-to nigiru”	The festival was extremely crowded, so Emily clutched her father’s hand very tightly to keep from getting lost.	more
devour	むさぼり食う “musabori-ku”	A: Toby, did you see that nature program about tigers last night? B: Yeah, I enjoyed it, but some scenes were too much for my young daughter. She got upset when she saw a tiger killing and devouring its prey.	more
discredit	評判を悪くする “hyoban-o waruku-suru”	Rumors that the presidential candidate accepted illegal campaign financing proved to be false. They were simply an attempt by his critics to discredit him.	less
enrollment	入学 “nyugaku”	Faced with the lowest student enrollment in 20 years, the university had no choice but to make a number of budget and staff cuts.	less
envision	心に描く “kokoro-ni egaku”	Some scientists say that the failure of governments to address global warming means it is difficult to envision a future without severe environmental damage.	less
extravagant	金遣いの荒い “kane-zukai-no arai”	The rock star’s simple lifestyle contrasted sharply with the extravagant habits of many others in his profession.	less
hypocritical	見せかけの “misekake-no”	A: Darling, did you talk to Sylvia about the importance of attending all her classes at college? B: I did, but I felt a bit hypocritical. After all, I often skipped classes when I was a student.	less
outgoing	社交的な “shakoteki-na”	Sara’s outgoing personality makes her an excellent sales manager. She enjoys meeting new people and is able to put strangers at ease within minutes.	more

Appendix 3.7

The Format of the Contextualized Recall Test Used in Experiment 2

先ほど学習した英単語を思い出して、英文中の空欄に最もよくあてはまるものを記入してください。

01	Some scientists say that the failure of governments to address global warming means it is difficult to （　　）a future without severe environmental damage.
02	A: Sergeant Holmes, it's important that we meet outside the suspect's house at precisely 10:05 p.m. tonight. B: I understand, sir. Let's （　　）our watches to make sure we have the same time.
03	The building of a huge shopping mall was so （　　）that it took years of negotiations with local residents before construction could begin.
04	The flight attendant tried to help the woman put her suitcase into the overhead compartment, but it was too （　　）to fit.
05	Although Bill's salary is （　　）for his day-to-day living costs, he finds it impossible to save any money for the future.
06	A: Toby, did you see that nature program about tigers last night? B: Yeah, I enjoyed it, but some scenes were too much for my young daughter. She got upset when she saw a tiger killing and （　　）its prey.
07	Rumors that the presidential candidate accepted illegal campaign financing proved to be false. They were simply an attempt by his critics to （　　）him.
08	Sara's （　　）personality makes her an excellent sales manager. She enjoys meeting new people and is able to put strangers at ease within minutes.
09	Faced with the lowest student （　　）in 20 years, the university had no choice but to make a number of budget and staff cuts.
10	The festival was extremely crowded, so Emily （　　）her father's hand very tightly to keep from getting lost.

指示があるまで次のページに進まないでください。

Appendix 3.8

The Target Words, Translations, and Contexts Used in Experiment 3

Target	Translation	Context	Imageability
		Decontextualized condition	
aisle	廊下 “roka”		N.A.
erupt	噴火する “funka-suru”		N.A.
folly	愚かさ “oroka-sa”		N.A.
grate	不快感を与える “fukai-kan o ataeru”		N.A.
lessen	少なくする “sukunaku-suru”		N.A.
limp	足を引きずって歩く “ashi o hikizutte aruku”		N.A.
orchid	ランの花 “ran no hana”		N.A.
proxy	代理 “dairi”		N.A.
reptile	は虫類 “hachurui”		N.A.
suffrage	選挙 “senkyo”		N.A.
		Contextualized+IMG condition	
crook	曲げる “mageru”	The baby tried to crook her cute fingers.	more
havoc	災害 “saigai”	Let’s go see Twister, a movie with a lot of havoc.	more
mar	台無しにする “dainashi ni suru”	The heavy rain marred our trip to Kyoto.	more
prosper	栄える “sakaeru”	Nintendo is a big, prospering business company.	more
reap	刈り取る “karitoru”	My uncle reaps a lot of flowers by hand.	more
rite	儀式のやり方 “gishiki no yarikata”	My sister’s wedding was conducted following traditional Japanese rites.	more

Appendix 3.8 (*continued*)

Target	Translation	Context	Imageability
saga	長編物語 “chohen monogatari”	You can read old sagas in the library.	more
saloon	広い部屋 “hiroi heya”	Let’s go to the saloon and drink tea.	more
tenure	在職期間 “zaishoku kikan”	The president said, “My tenure is three years.”	more
womb	内部 “naibu”	The teacher said the earth’s womb is 5000° C.	more
		Contextualized–IMG condition	
adorn	装飾する “soshoku-suru”	Some of them will adorn the wall tomorrow.	less
despise	ひどく嫌う “hidoku kirau”	One of us said we shouldn’t despise them.	less
glacier	氷河 “hyoga”	She thought about the glacier for a moment.	less
grievance	不満 “fuman”	He wants to take his grievance to them.	less
herring	魚のニシン “sakana no nishin”	They talked about herrings a few weeks ago.	less
pebble	小石 “koishi”	They thought somebody had carried the pebbles there.	less
perish	死ぬ “shinu”	I heard that they had perished in battle.	less
stalk	大またに歩く “omatani aruku”	One day, she saw it stalking the street.	less
wrath	激しい怒り “hageshi ikari”	He remembered that they were in quiet wrath.	less
zeal	熱意 “netsui”	I think they prepared for it with zeal.	less

Appendix 3.9

The Format of the Translation Test Used in Experiment 3

■英単語の意味を日本語で書きなさい。

1	mar	(1)
2	crook	(2)
3	womb	(3)
4	prosper	(4)
5	reap	(5)
6	tenure	(6)
7	saloon	(7)
8	havoc	(8)
9	rite	(9)
10	saga	(10)

Appendix 3.10

The Format of the Contextualized Recall Test Used in Experiment 3

■文脈中の（　　）に入る、最も適切な英単語を書きなさい。

1	The heavy rain (　　) our trip to Kyoto.	(1)
2	Let's go to the (　　) and drink tea.	(2)
3	Nintendo is a big, (　　) business company.	(3)
4	You can read old (　　) in the library.	(4)
5	The baby tried to (　　) her cute fingers.	(5)
6	My uncle (　　) a lot of flowers by hand.	(6)
7	Let's go see *Twister,* a movie with a lot of (　　).	(7)
8	The president said, "My (　　) is three years."	(8)
9	The teacher said the earth's (　　) is 5000° C.	(9)
10	My sister's wedding was conducted following traditional Japanese (　　).	(10)

Appendix 4.1

The Questionnaire Format Used in Experiment 4 (in Japanese)

あなたは普段、どのような方法で英単語を学習していますか？各文について、以下の要領でどれくらいあてはまるかを 5 段階で評価して、あてはまる数字に丸をつけてください。全部で 25 項目あります。

全くあてはまらない	あまりあてはまらない	どちらとも言えない	ややあてはまる	完全にあてはまる
1	2	3	4	5

1. 覚えようとする英単語を定期的に復習する。

1　2　3　4　5

2. いつでも学習できるように英単語の本やリストを持ち歩く。

1　2　3　4　5

3. 一定期間内で何個かの英単語を覚えることにしている（「一日に 10 語」など）。

1　2　3　4　5

4. 授業で学習したものに加えて、それ以外の英単語も学習しようとする。

1　2　3　4　5

5. 英単語の学習のための時間を作るようにする。

1　2　3　4　5

6. テスト対策（TOEIC、TOEFL、英検などの対策）として英単語を学習する。

1　2　3　4　5

7. 英単語の学習・確認・復習のために、自分自身のやり方を用いる。

1　2　3　4　5

8. リーディングやリスニングを多くすることで、英単語にふれる機会を増やす。

1　2　3　4　5

9. 英単語にふれる機会を増やすための学習環境を整えるようにする。

1　2　3　4　5

10. 英単語の学習のためにメディアを活用する（テレビ、ラジオ、インターネット、携帯電話、映画など）。

1　2　3　4　5

11. 英単語を使うということを前提に学習する。

1　2　3　4　5

12. 英単語を学習するときに、その意味に関係するイメージを心の中に思い浮かべる。

1　2　3　4　5

13. 英単語を学習するときに、自分の体験を関連づける。

1　2　3　4　5

14. 英単語を学習するときに、スペリングや単語の形に関するイメージを心の中に思い描く。

1　2　3　4　5

15. 英単語を学習するときに、語呂合わせを使う（ダジャレで覚える、キーワード法を使う）。

1　2　3　4　5

16. 英単語を学習するときに、その意味が良い意味か悪い意味かを思い浮かべる。

1　2　3　4　5

Appendix 4.1 (*continued*)

17. 英単語を何度も繰り返し書いて学習する。

1　　2　　3　　4　　5

18. 英単語をノートやカードに書いて学習する。

1　　2　　3　　4　　5

19. 英単語のスペリングだけではなく、意味も書いて学習する。

1　　2　　3　　4　　5

20. 英単語を何度も繰り返し音読して学習する。

1　　2　　3　　4　　5

21. 英単語の意味だけでなく、発音を覚えるために音読して学習する。

1　　2　　3　　4　　5

22. 英単語を覚えるときに例文を音読する。

1　　2　　3　　4　　5

23. すでに知っている同義語（begin と start など）や反義語（positive と negative など）と関連付けて、英単語を学習する。

1　　2　　3　　4　　5

24. 覚えようとする英単語と一緒に、同義語や反義語を学習する。

1　　2　　3　　4　　5

25. 覚えようとする英単語と一緒に、意味、発音、形などが似ていたり、関連性のある単語をまとめて学習する。

1　　2　　3　　4　　5

※先ほどは、どのような学習方法で英単語を覚えましたか？簡潔に記入してください。

（　　）

指示があるまで次のページに進まないでください。

Note. The 25 items using a 5-point Likert scale were adapted from Mizumoto and Takeuchi (2009); each statement was translated by the present author. The English version is presented in Appendix 4.2. The additional item that asked participants to describe their own learning methods was used for the learner-type analysis.

Appendix 4.2

The Results of Mizumoto and Takeuchi's (2009) Questionnaire

Description	Selected frequency (N = 23)				
	1	*2*	*3*	*4*	*5*
Self-management					
1. I regularly review the vocabulary I learned to check if I remember it.	0	5	4	11	3
2. I keep a vocabulary book or word list to check the vocabulary anytime I wish.	4	6	5	5	3
3. I try to make it a rule to memorize a certain number of words in a specific time period.	9	7	2	4	1
4. I try to learn extra vocabulary in addition to what I am taught in class.	2	7	4	8	2
5. I try to take time for vocabulary learning.	1	10	3	6	3
6. I consciously set aside time to study vocabulary in order to prepare for tests.	2	4	3	7	7
7. I use my own methods for remembering, checking, or reviewing vocabulary.	0	0	8	9	6
Subtotal (Items 1–7)	18 (11.2%)	39 (24.2%)	29 (18.0%)	50 (31.1%)	25 (15.5%)
Input-seeking					
8. I try to expose myself to English vocabulary by reading or listening a lot.	1	6	5	9	2
9. I try to manage the learning environment so as to expose myself to English vocabulary.	1	6	8	7	1
10. I try to make use of the media to learn vocabulary.	4	7	5	4	3
11. I study vocabulary with the intention of using it.	1	6	6	9	1
Subtotal (Items 8–11)	7 (7.6%)	25 (27.2%)	24 (26.1%)	29 (31.5%)	7 (7.6%)
Imagery					
12. When I try to remember vocabulary, I make a mental picture of what can be associated with a word's meaning.	0	4	3	10	6
13. When I try to remember vocabulary, I link my personal experiences to it.	5	6	2	8	2
14. When I try to remember vocabulary, I create an image of the spellings or orthographic forms.	1	3	3	9	7

Appendix 4.2 (*continued*)

Description	Selected frequency (N = 23)				
	1	*2*	*3*	*4*	*5*
15. When I try to remember vocabulary, I use the keyword method.	6	6	2	7	2
16. When I try to remember vocabulary, I imagine whether the meaning of the word is negative or positive.	0	3	3	11	6
Subtotal (Items 12–16)	12 (10.4%)	22 (19.1%)	13 (11.3%)	45 (39.1%)	23 (20.0%)
Writing Rehearsal					
17. When I try to remember vocabulary, I write it repeatedly.	3	5	3	6	6
18. When I try to remember vocabulary, I write it on a note or a card.	3	6	2	10	2
19. When I try to remember vocabulary, I remember not only the meaning but also the spelling of the word by writing it.	4	5	2	9	3
Subtotal (Items 17–19)	10 (14.5%)	16 (23.2%)	7 (10.1%)	25 (36.2%)	11 (15.9%)
Oral Rehearsal					
20. When I try to remember vocabulary, I say it aloud repeatedly.	1	1	6	7	8
21. When I try to remember vocabulary, I vocalize it to remember not only the meaning but also the pronunciation of the word.	1	3	4	7	8
22. When I try to remember vocabulary, I say the sample sentence aloud.	2	3	6	6	6
Subtotal (20–22)	4 (5.8%)	7 (10.1%)	16 (23.2%)	20 (29.0%)	22 (31.9%)
Association					
23. When I try to remember vocabulary, I associate it with the synonyms or antonyms I already know.	0	5	3	8	7
24. When I try to remember vocabulary, I also memorize the synonyms or antonyms of the word.	0	7	3	11	2
25. When I try to remember vocabulary, I memorize words similar to it or the related words in a group.	0	8	4	8	3
Subtotal (23–25)	0 (0.0%)	20 (29.0%)	10 (14.5%)	27 (39.1%)	12 (17.4%)

Appendix 4.3

The Results of the Questionnaire on Learning Styles (in Japanese)

No.	Response
	Word-oriented
1.	単語のスペルとその意味を中心に学習した。口を動かしたりもした。
2.	hypo-とか extra-とか、単語それぞれに前がついていることを暗記して、それから中身を覚えた。
3.	日本語をもので隠して訳を想像する。スペルを書く風にして覚える。何周も繰り返す。
4.	スペルと発音が結びつきにくいものは書いて、他は黙読していた。
5.	パッと見の感じをつかんで、心の中で繰り返し音読。意味はあっさり。
6.	ペンで空中で書く、単語を分解して意味を考える（extra-, contra-）。
	Context-oriented
1.	例文から英単語とその意味を覚えようとした。
2.	英単語の意味を見て、例文を読み、イメージ化した。
3.	例文の意味を考えて読むようにし、語呂合わせも用いた。
4.	知っている単語と関連づけたり、文脈の中でどのような単語と用いられているかを覚えるようにした。また、単語から連想できるイメージと意味を結びつけるようにした。
5.	日本語の意味の部分を隠して覚えました。
6.	一通り全部の文章を読んでから単語の意味だけを何回もくり返し見た。
	Word-and-context-oriented
1.	実際に何度かスペルを書いてから例文に目を通しました。
2.	単語のスペルと意味と例文を交互に見て学習した。
3.	単語を数回音読後、意味を隠して確認。その後例文を読む。
4.	例文で意味を覚え、英語だけを見て日本語訳が言えるようになるよう繰り返した。
5.	初めは頭の中で暗誦、次に 1～2 回書きとり、例文で印象付け。
6.	音節ごとに分解、例文の黙読、単語を書く
7.	スペル→発音を頭の中でしてみる、発音からスペルをイメージ。意味→文脈と関連付けて。
8.	単語と意味をくり返し書く→書きながら自分のイメージとむすぶ→例文を読む。
9.	単語を見て意味がすぐ浮かぶように、「英単語」と「意味」を繰り返しみる。例文を書く。
	Others
1.	スペリングの形と、例文で文脈の前後の部分だけ覚えた。
2.	繰り返し書き、音読もくり返し行った。例文は、目標語の前後の語だけ覚えた。

Note. The two participants were excluded from analysis because they focused on some specific words in the context sentences (labeled as “others”).

Appendix 4.4

The Target Words, Translations, and Contexts Used in Experiment 5

Target	Translation	Context	Imageability
indite	執筆する	Ann bought a sheet of paper to indite a long letter.	more
	“shippitsu-suru”	Then they asked me, “Why don’t you indite another one?”	less
reminisce	思い出にふける	Mike doesn’t reminisce about the Christmas day with his girlfriend.	more
	“omoide ni fukeru”	Maybe it happened, but she did not reminisce about it.	less
gloaming	夕暮れ時	We do not have a long gloaming time in Australia.	more
	“yugure-doki”	They danced in the gloaming and talked with each other.	less
repast	食事	Mary and Paul had a repast together in the Italian restaurant.	more
	“shokuji”	They found the person in the garden after the repast.	less
tram	路面電車	The tram left the station and went into the dark.	more
	“romen densha”	Everyone on the tram and the ship laughed at him.	less
sanatorium	療養所	Kate is a new doctor at the sanatorium in London.	more
	“ryoyo-jo”	She read the paper during her time spent in the sanatorium.	less
visage	顔つき	When the boy smiled, his visage looked happy and excited.	more
	“kao-tsuki”	They stood up and saw their visage in the mirror.	less
esplanade	遊歩道	All the children played in the narrow esplanade last Sunday.	more
	“yuhodo”	They thought, “We should go back through the esplanade.”	less
don	着用する	The little girl tried to don a large white coat.	more
	“chakuyo-suru”	I don’t know anything about art, so I don’t don expensive clothes.	less
mosey	ぶらりと訪れる	I want to come to Paris and mosey along with you.	more
	“burari-to otozureru”	They wanted to mosey over to you, but they did not.	less

Appendix 4.5

The Format of the First Contextualized Recall Test Used in Experiment 5 (Contextualized Test B)

■先ほど学習した英単語を思い出して、英文中の空欄に最もあてはまるものを記入してください。【5 分】

	英文	解答欄
01	We do not have a long（ ） time in Australia.	（　　）
02	Mike doesn't（ ） about the Christmas day with his girlfriend.	（　　）
03	Kate is a new doctor at the（ ） in London.	（　　）
04	The little girl tried to（ ） a large white coat.	（　　）
05	When the boy smiled, his（ ） looked happy and excited.	（　　）
06	I want to come to Paris and（ ） along with you.	（　　）
07	Mary and Paul had a（ ） together in the Italian restaurant.	（　　）
08	Ann bought a sheet of paper to（ ） a long letter.	（　　）
09	All the children played in the narrow（ ） last Sunday.	（　　）
10	The（ ） left the station and went into the dark.	（　　）

Appendix 4.6

The Format of the Translation Test Used in Experiment 5 (Decontextualized Test A)

■先ほど学習した英単語の意味を思い出して、日本語で記入してください。【3 分】

	英単語	解答欄
01	don	意味　（　　）
02	esplanade	意味　（　　）
03	gloaming	意味　（　　）
04	indite	意味　（　　）
05	mosey	意味　（　　）
06	reminisce	意味　（　　）
07	repast	意味　（　　）
08	sanatorium	意味　（　　）
09	tram	意味　（　　）
10	visage	意味　（　　）

Appendix 4.7

The Format of the Recognition Test Used in Experiment 5 (Decontextualized Test B)

■先ほど学習した英単語の意味を１つ選んで、選択肢 1～4 の番号に丸をつけてください。【2 分】

英単語		選択肢
01	don	1. 執筆する　2. 着用する　3. 思い出にふける　4. ぶらりと訪ねる
02	esplanade	1. 食事　2. 遊歩道　3. 路面電車　4. 顔つき
03	gloaming	1. 夕暮れ時　2. 療養所　3. 顔つき　4. 遊歩道
04	indite	1. ぶらりと訪ねる　2. 着用する　3. 執筆する　4. 思い出にふける
05	mosey	1. 着用する　2. 思い出にふける　3. ぶらりと訪ねる　4. 執筆する
06	reminisce	1. 思い出にふける　2. 執筆する　3. 着用する　4. ぶらりと訪ねる
07	repast	1. 療養所　2. 夕暮れ時　3. 食事　4. 路面電車
08	sanatorium	1. 路面電車　2. 食事　3. 遊歩道　4. 療養所
09	tram	1. 顔つき　2. 路面電車　3. 夕暮れ時　4. 遊歩道
10	visage	1. 夕暮れ時　2. 顔つき　3. 療養所　4. 食事

Appendix 4.8

The Format of the Second Contextualized Recall Test Used in Experiment 5 (Contextualized Test A)

■先ほど学習した英単語を思い出して、英文中の空欄に最もあてはまるものを記入してください。【5 分】

■ただし、先ほど読んだ英文とは異なるものが書かれていますので、注意してください。

英文		解答欄
01	In the（ ）, Lucy hurried to the airport by taxi.	(　　)
02	Many people will（ ）about him for a long time.	(　　)
03	They thought, “The beds in the（ ）are for sick people.”	(　　)
04	They really liked to（ ）old gloves on their hands.	(　　)
05	I could not see his（ ）or his body at that time.	(　　)
06	Jane decided to（ ）over to her old grandfather every morning.	(　　)
07	He had a（ ）and tea with Sue in the kitchen.	(　　)
08	He is young, but he can（ ）it by himself.	(　　)
09	John met his friend at the corner of the（ ）.	(　　)
10	The（ ）began to slow down near that area.	(　　)

Appendix 5.1

The Format of the Mental Imagery Rating Task Used in Experiment 6

・これが本日最後の課題です。先ほど、英単語を学習した時のことを思い出してください。

・例文を読んだとき、あなたはその内容（記述された状況）を心の中で思い浮かべていましたか？

・それぞれの例文の内容をイメージしたかどうかについて、当てはまるものに〇をつけてください。

01	apprentice 見習い	Before John became a professional chef, he spent many years working as an **apprentice** to a master chef at a famous restaurant in London.
	イメージしなかった　部分的にイメージした　イメージした　その他(　　)	
02	discredit ～の評判を悪くする	Rumors that the presidential candidate accepted illegal campaign financing proved to be false. They were simply an attempt by his critics to **discredit** him.
	イメージしなかった　部分的にイメージした　イメージした　その他(　　)	
03	assorted 各種取りそろえた	A: Darling, what kind of cheese should we serve at our dinner party on Saturday? B: I'm going to prepare a tray with **assorted** cheeses on it. That way, people can choose what they like.
	イメージしなかった　部分的にイメージした　イメージした　その他(　　)	
04	enrollment 入学	Faced with the lowest student **enrollment** in 20 years, the university had no choice but to make a number of budget and staff cuts.
	イメージしなかった　部分的にイメージした　イメージした　その他(　　)	
05	clutch ～をぐいと握る	The festival was extremely crowded, so Emily **clutched** her father's hand very tightly to keep from getting lost.
	イメージしなかった　部分的にイメージした　イメージした　その他(　　)	
06	envision ～を心に描く	Some scientists say that the failure of governments to address global warming means it is difficult to **envision** a future without severe environmental damage.
	イメージしなかった　部分的にイメージした　イメージした　その他(　　)	
07	devour ～をむさぼり食う	A: Toby, did you see that nature program about tigers last night? B: Yeah, I enjoyed it, but some scenes were too much for my young daughter. She got upset when she saw a tiger killing and **devouring** its prey.
	イメージしなかった　部分的にイメージした　イメージした　その他(　　)	
08	extravagant 金遣いの荒い	The rock star's simple lifestyle contrasted sharply with the **extravagant** habits of many others in his profession.
	イメージしなかった　部分的にイメージした　イメージした　その他(　　)	
09	outgoing 社交的な	Sara's **outgoing** personality makes her an excellent sales manager. She enjoys meeting new people and is able to put strangers at ease within minutes.
	イメージしなかった　部分的にイメージした　イメージした　その他(　　)	
10	hypocritical 見せかけの	A: Darling, did you talk to Sylvia about the importance of attending all her classes at college? B: I did, but I felt a bit **hypocritical**. After all, I often skipped classes when I was a student.
	イメージしなかった　部分的にイメージした　イメージした　その他(　　)	

※ひと通り回答が終わったら、「終わりました」と声を掛けてください。

Appendix 5.2

Summary of Think-Aloud Protocols Obtained for the More and for the Less Imageable Items in Experiment 6

	More imageable						Less imageable							
	Word-based		Context-based			Monitoring	Word-based		Context-based			Monitoring		
	L2	L1	Analytic		Inferential		L1	L2	Analytic		Inferential			
	AEF	BG	CDH	(CH)	IJ	K	AEF	BG	CDH	(CH)	IJ	K	Other	Total
S1	35	34	27	(17)	0	0	24	24	25	20	0	1	3	173
S2	26	51	8	(8)	0	0	27	47	6	6	0	3	14	182
S3	3	7	12	(10)	0	3	4	5	23	18	0	0	10	67
S4	21	4	1	(1)	0	2	23	4	1	1	0	0	4	60
S5	8	8	33	(25)	0	0	11	8	48	36	0	0	4	120
S6	16	9	14	(13)	0	0	12	8	20	18	0	0	3	82
S7	29	5	30	(29)	7	19	46	8	50	44	1	23	13	231
S8	33	12	16	(12)	7	2	36	13	32	29	3	4	9	167
S9	13	3	16	(13)	0	0	10	0	17	15	0	0	3	62
Total	184	133	157	128	14	26	193	117	222	187	4	31	63	1,144

Note. The 11 strategy categories from A to K were sorted into word-based, context-based, and self-monitoring strategies. For example, the aggregate subtotal for word form (A), translation (B), word analysis (E), L2 word analogy (F), and L1 word analogy (G) was compared to the subtotal of local context (C), global context (D), sentence analysis (H), implicit idea (I), and imagery report (J) strategies. The frequency of self-monitoring (K) and the other strategies were excluded from this analysis.

Appendix 5.3

Summary of Mental Imagery Ratings and Time Allocation Data in Experiment 6

	Mental imagery ratings			Time allocation		
	More	Less	All	More	Less	All
S1	0.40	0.20	0.30	136	203	433
S2	1.00	0.70	0.85	298	624	922
S3	0.90	0.50	0.70	208	264	592
S4	0.50	0.50	0.50	341	455	796
S5	0.90	0.80	0.85	275	358	640
S6	0.80	0.50	0.65	71	116	474
S7	0.70	0.80	0.75	105	181	338
S8	0.40	0.40	0.40	101	103	472
S9	0.70	0.60	0.65	323	370	694
M	0.70	0.56	0.63	206.44	297.11	595.67
(*SD*)	(0.22)	(0.19)	(0.19)	(105.77)	(171.29)	(187.49)

Note. The mental imagery ratings were based on each participant's report of whether or not she or he called up mental imagery from the context in question during learning (1 = called up imagery; 0 = did not). The time allocation data show how many seconds each participant spent on items with more and less imageable contexts respectively and on the whole word list. The total amount of time, displayed in the "All" column, is not equal to the sum of "More" and "Less" because it include a participant's thinking time unrelated to learning of a specific item (e.g., when a participant counted how many items were in the word list).

Appendix 6.1

The Results of the Latent Semantic Analysis (LSA) and the Keyword Abstraction

Exp.	Target word	Context	LSA
		More imageable	
1	train	The train left the station and rushed into the dark.	.68
1	street	At the corner of the street she met Laurie.	.65
5	don	The little girl tried to don a large white coat.	.65
1	hospital	I am a doctor at the London Hospital.	.61
1	visit	I visited him every day, and talked to him.	.48
2, 4, 6	apprentice	Before John became a professional chef, he spent many years working as an apprentice to a master chef at a famous restaurant in London.	.46
1	write	Your brother wrote his name on this paper.	.45
3	tenure	The president said, "My tenure is three years."	.39
3	crook	The baby tried to crook her cute fingers.	.33
3	saloon	Let's go to the saloon and drink tea.	.32
5	sanatorium	Kate is a new doctor at the sanatorium in London.	.30
2, 4, 6	outgoing	Sara's outgoing personality makes her an excellent sales manager. She enjoys meeting new people and is able to put strangers at ease within minutes.	.26
2, 4, 6	assorted	Darling, what kind of cheese should we serve at our dinner party on Saturday? I'm going to prepare a tray with assorted cheeses on it. That way, people can choose what they like.	.25
5	tram	The tram left the station and went into the dark.	.25
3	mar	The heavy rain marred our trip to Kyoto.	.24
5	visage	When the boy smiled, his visage looked happy and excited.	.22
2, 4, 6	devour	Toby, did you see that nature program about tigers last night? Yeah, I enjoyed it, but some scenes were too much for my young daughter. She got upset when she saw a tiger killing and devouring its prey.	.21
3	rite	My sister's wedding was conducted following traditional Japanese rites.	.21
5	mosey	I want to come to Paris and mosey along with you.	.17

Appendix 6.1 (*continued*)

Exp.	Target word	Context	LSA
3	havoc	Let's go see Twister, a movie with a lot of havoc.	.14
3	saga	You can read old sagas in the library.	.14
3	prosper	Nintendo is a big, prospering business company.	.13
3	womb	The teacher said the earth's womb is 5000 °C.	.09
2, 4, 6	clutch	The festival was extremely crowded, so Emily clutched her father's hand very tightly to keep from getting lost.	.08
3	reap	My uncle reaps a lot of flowers by hand.	.08
5	gloaming	We do not have a long gloaming time in Australia.	.06
5	repast	Mary and Paul had a repast together in the Italian restaurant.	.04
		Less imageable	
1	face	As she stood up to see if her bag was all right, she saw her face in the mirror white, with big, round eyes.	.69
1	evening	They danced together many times that evening, and Archie told her his plans.	.65
1	remember	Perhaps this happened. Or perhaps she's ill and can't remember who she is.	.61
1	lunch	Paul found the Minister sitting alone in the garden after lunch, smoking a large cigar, his big red hands folded over his large stomach, a soft hat over his eyes.	.43
1	wear	I don't know anything about art, and I haven't met any grand people, and I don't wear expensive clothes, but up to now, none of that has worried me.	.39
2, 4, 6	enrollment	Faced with the lowest student enrollment in 20 years, the university had no choice but to make a number of budget and staff cuts.	.35
3	pebble	They thought somebody had carried the pebbles there.	.32
3	wrath	He remembered that they were in quiet wrath.	.31
3	zeal	I think they prepared for it with zeal.	.31
2, 4, 6	discredit	Rumors that the presidential candidate accepted illegal campaign financing proved to be false. They were simply an attempt by his critics to discredit him.	.30
3	perish	I heard that they had perished in battle.	.30
2, 4, 6	extravagant	The rock star's simple lifestyle contrasted sharply with the extravagant habits of many others in his profession.	.25

Appendix 6.1 (*continued*)

Exp.	Target word	Context	LSA
5	sanatorium	She read the paper during her time spent in the sanatorium.	.22
2, 4, 6	envision	Some scientists say that the failure of governments to address global warming means it is difficult to envision a future without severe environmental damage.	.21
3	despise	One of us said we shouldn't despise them.	.21
3	glacier	She thought about the glacier for a moment.	.21
5	visage	They stood up and saw their visage in the mirror.	.18
3	grievance	He wants to take his grievance to them.	.16
5	tram	Everyone on the tram and the ship laughed at him.	.16
3	adorn	Some of them will adorn the wall tomorrow.	.15
2, 4, 6	hypocritical	Darling, did you talk to Sylvia about the importance of attending all her classes at college? I did, but I felt a bit hypocritical. After all, I often skipped classes when I was a student.	.14
5	gloaming	They danced in the gloaming and talked with each other.	.14
3	herring	They talked about herrings a few weeks ago.	.12
5	repast	They found the person in the garden after the repast.	.12
5	mosey	They wanted to mosey over to you, but they did not.	.12
3	stalk	One day, she saw it stalking the street.	.10

Note. The semantic overlap of each target word and its context was calculated using the LSA tool. The top 20 keywords of the more and less imageable contexts are underlined. The less imageable context for *don* was removed from the table. The 10 pseudo words were replaced with the following synonyms: *ancon* (hospital), *cader* (lunch), *dangy* (street), *denent* (remember), *faddam* (write), *hodet* (face), *masco* (train), *pacon* (wear), *sagod* (visit), and *tasper* (evening). Among the 30 items *indite*, *reminisce*, and *esplanade* were not analyzed, since they were not registered in the LSA corpus.