Processing and Interpreting Unknown Words With Morphological and Contextual Information Among Japanese EFL Learners: Focusing on the Semantic Transparency of Morphemes and Learner Proficiency

A Dissertation
Submitted to the University of Tsukuba
In Partial Fulfillment of the Requirements for
the Degree of
Doctor of Philosophy in Linguistics

Kentaro SUZUKI

2018

Abstract of the Dissertation

Processing and Interpreting Unknown Words With Morphological and Contextual Information Among Japanese EFL Learners: Focusing on the Semantic Transparency of Morphemes and Learner Proficiency

by

Kentaro SUZUKI

Having knowledge of the vocabulary used in a text enables readers to understand the ideas described in the text; however, unknown words interfere with accurate and fluent reading. To fill the lexical gap in a text, readers often make lexical inferences by employing the linguistic and nonlinguistic clues provided by the text and combining them with the readers' general knowledge of the world (e.g., Haastrup, 1991).

For lexical inferencing, morphemes of an unknown word and context are useful sources of information (Wesche & Paribakht, 2010). Further, morphemes of a word convey concrete semantics inherent to the word (Nagy & Anderson, 1984), and semantically transparent words such as *sunlight* can be inferred from morphological information. However, in some cases, a morpheme-based meaning is not related to the actual meaning of the word, for instance, the opaque word *honeymoon*. Earlier studies have revealed that this type of readers' overreliance on word-based information is a very frequent mistake (Bensoussan & Laufer, 1984; Laufer, 1989), which makes lexical inferencing problematic, particularly for second language (L2) learners. Therefore, successful lexical inferencing requires one to consider contextual information, as well, since it limits the possible meanings of a word in a specific scenario.

To date, L2 studies have revealed the challenges of using two sources of information,

especially for less proficient learners (Hamada, 2014; Mori, 2002; Mori & Nagy, 1999). The potential difficulties in inferring opaque unknown words are caused by multiple processes, such as perceiving semantic inconsistencies between morphemes and context and, thereafter, generating contextually appropriate meanings (Oakhill, Cain, & Nesi, 2016).

However, few studies have examined the process involved in the combined use of the two sources, which makes it difficult to identify the particular challenges and possible educational interventions for learners. Therefore, this study examines the processing and interpretation of semantically transparent and opaque unknown words by Japanese learners of English as a foreign language (EFL). Due to the critical role of learner proficiency in the use of contextual information (e.g., Hamada, 2014; Mori, 2002), this study targets learners of different proficiency (intermediate and beginner) levels and examines various hitherto unresolved issues in each proficiency group by conducting two experimental studies (Study 1 and 2).

Study 1 comprised three experiments (Experiments 1–3), targeted intermediate-level Japanese EFL learners, and examined the processing and interpretation of unknown words based on morphological and contextual information and resulting text comprehension. First, to identify the source of difficulties involved in appropriate inferences, Experiment 1 examined whether Japanese EFL learners could perceive the semantic relationship (in/consistency) between morphemes and context and appropriately interpret unknown words according to the semantic transparency of words. In addition, the experiment checked their sentence representations. Participants read sentential context with target compound words: their literal interpretation of morphemes was either consistent (transparent) or inconsistent (opaque) with the contextual meaning in a word-by-word self-paced reading task. Subsequently, they translated the context into Japanese. The results revealed that although the participants were sensitive to the semantic relationship between the two sources even in neutral contexts, they interpreted opaque words literally in many cases. This suggests that the difficulty in inferencing

unknown words, particularly opaque ones, arises from generating contextually appropriate meanings. In addition, participants sometimes distorted interpretations of contextual information.

Experiment 2 examined the relationship between text processing and semantic memory representation of unknown words constructed from reading. The participants read two-sentence passages sentence by sentence and, immediately afterward, answered a comprehension question that queried the meaning of the target words immediately after reading. The results showed that participants attempted to infer the meanings of unknown words even after reading the sentence that introduced the target words. However, this attempt did not always lead to appropriate interpretation in many cases.

Experiment 3 examined the relationship between the use of a lexical inference strategy and interpretation of unknown words. Participants inferred target words while verbalizing their thoughts (think-aloud task). The results showed that, in general, participants provided contextually appropriate interpretation of opaque words by using various metacognitive strategies and linguistic clues. The better inference outcomes relative to the prior experiments were attributed to the high importance of unknown words in a lexical inference task than in normal reading. In addition, the qualitative analysis of literal interpretations of opaque words revealed that such interpretations were sometimes caused by changes in interpretation of the context, which reflects the readers' mental effort to meaningfully interpret the sentence.

Study 2 comprised two experiments (Experiments 4 and 5), focused on beginner-level learners who experienced particular difficulty in using the two sources, and examined ways to support their successful inferencing. Experiment 4 examined whether participants could understand the semantic relationship between the two sources, as well as their ability to use contextual information for inferences, to determine the focus of support in Experiment 5. Participants were explicitly asked to judge the semantic in/consistency between morphemes

and context, both with and without time pressure. The results revealed that they could generally understand the relationship and benefit more from informative contexts for activating the semantics of upcoming words than from neutral contexts.

Experiment 5 examined the effects of onetime inference training on participants' lexical inference performance. The participants, who were divided into Upper, Middle, and Lower groups, practiced inferencing the meaning of unknown words using an inference worksheet, which described when and how to use morphological and contextual information, and took pre- and post-inference tests, as well. A comparison between the pre-test and training revealed that training reduced the inappropriate literal and morpheme-based interpretation of opaque words (in all groups). However, a slight increase in appropriate interpretation was observed for the Middle and Upper groups, but was not confirmed for the Lower group. In addition, a comparison between the pre- and post-tests showed that the aforementioned effects of training were not retained one week after the post-test.

The findings from the five experiments revealed the characteristics of use of morphological and contextual information in interpreting unknown words based on learner proficiency. Intermediate-level learners have sufficient linguistic and metacognitive skills to use morphological and contextual information according to the semantic transparency of morphemes. However, beginner-level learners experience more difficulty in both understanding the semantic relationship between morphemes and context and subsequent generation of contextually appropriate meaning due to their inaccurate contextual understanding and their limited cognitive skills in performing the multiple processes involved in successful inferences.

These findings have pedagogical implications for educators on how they can help their students interpret unknown words in texts by considering the proficiency-related characteristics. Despite having some limitations, the current study is significant since it provides valuable insight into the processes of EFL lexical inferencing and the role of learner proficiency.

Acknowledgements

I am profoundly thankful to many people for their invaluable support while completing my studies at the Graduate School of Humanities and Social Sciences at the University of Tsukuba. I would like to extend my heartfelt gratitude especially to the following people.

First and foremost, I would like to express my deepest appreciation to my supervisor, Professor Yuji Ushiro at the University of Tsukuba, who awakened my interest in L2 research. He always gives me beneficial and inspiring comments from both theoretical and educational viewpoints. His own experience of writing his dissertation while holding a job strongly motivated me to pursue my doctoral degree and also supported me whenever I faced challenges.

I would like to express my gratitude to Professor Hirosada Iwasaki at the University of Tsukuba. His deep knowledge of English vocabulary enabled me to clarify the standing position of this research. Also, what I learned from his classes on academic writing was very informative in writing this dissertation. I would like to thank Professor Akira Kubota at the University of Tsukuba. His meticulous comments helped me interpret the results more objectively. In addition, he gave me writing advice about how to convey the message of this dissertation more effectively and make it easier to understand. I am also deeply indebted to my external supervisor, Professor Shuhei Kadota at the Kwansei Gakuin University. His beneficial comments based on his deep insight into the field of reading and vocabulary research greatly improved the earlier version of this dissertation and also suggested the future direction of my research.

I am indebted to all the colleagues in Professor Ushiro's seminar for their invaluable advice, especially the following people. First, I would like to thank Dr. Shingo Nahatame, Dr. Yusuke Hasegawa, Dr. Yukino Kimura, and Dr. Akira Hamada for their helpful guidance during my studies in the master program. They taught me how to plan and conduct research, and their deep knowledge into reading and vocabulary research enabled me to design the current research

plan and discuss the findings from diverse perspectives. In addition, I would like to extend my gratitude to Dr. Natsumi Tanaka, Dr. Masaya Hosoda, and Dr. Yoshinobu Mori for their helpful support, especially during my doctoral course. They gave me critical and constructive comments on my research design. Also, I would like to thank Mr. Kozo Kamimura for carefully reviewing this manuscript and giving me useful suggestions.

My appreciation also extends to faculty members of Kyoei University. Their understanding and consideration of my research allowed me to complete this dissertation while working there. My teaching experience there and communication with my students inform me of difficulties that Japanese EFL learners have in reading comprehension at the lexical level, which became the main topic of this dissertation. I hope the implications of this research will help them learn English in the future.

Last but not least, I deeply thank my family. They always believe in me and warmly encouraged me. Thanks to them I was able to concentrate on my research. I could not have completed my dissertation without their help.

Kentaro SUZUKI

Table of Contents

Chapter 3

Study 1 Use of Morphological and Contextual Information in the Processing and Interpretation of Unknown Words and Their Effects on Text Comprehension by Intermediate-Level Japanese EFL Learners

Chapter 4
Study 2 Use of Morphological and Contextual Information in Lexical Inferencing by Beginner-Level Learners

Chapter 5 General Discussion

Chapter 6 Conclusion

List of Tables

Chapter 2

Chapter 3

List of Figures

Chapter 5

List of Appendices

Chapter 1

Introduction

1.1 Background of the Current Research

Vocabulary knowledge is an important component of both receptive and productive language use (e.g., Nation, 2013; Webb & Nation, 2017). In terms of reading comprehension, knowing many of the words in a given text makes it easy to understand the ideas they describe, while the presence of unfamiliar words can interfere with an accurate and fluent reading process. However, learners of a second language (L2) or English as a foreign language (EFL) are more likely to encounter unknown words while reading due to their relatively smaller vocabulary size in comparison to their first language (L1) counterparts. Therefore, learning to deal with unknown words is an important aspect of reading comprehension for L2 readers.

When readers encounter an unknown word in a text and want to know its meaning, they most often attempt to infer its meaning using available linguistic clues and combine these with their prior knowledge (Fraser, 1999; Paribakht & Wesche, 1999), which is called *lexical inferencing* (Haastrup, 1991). If successful, lexical inferencing leads to better text comprehension and fluent reading (Wesche & Paribakht, 2010). Additionally, active engagement with unknown words while inferencing contributes to incremental lexical development through reading (e.g., Huckin & Coady, 1999; Paribakht & Wesche, 1999).

While a variety of potential linguistic clues are available for inferencing, clues within the unknown word itself (e.g., word morphemes) and the sentence (e.g., sentence meaning) are most frequently used because of their high utility and availability (Wesche & Paribakht, 2010). Morphological information often includes semantics inherent to the word meanings, and many words are inferable by analyzing their morphological structures (Nagy & Anderson, 1984). For example, one may deduce the exact meaning of *sunlight* by decomposing it into morphemes

(i.e., *sun*, *light*) because it has semantically transparent morphological structures (word morphemes clearly contribute to its whole word meanings). However, the same strategy does not produce fruitful outcomes for words with semantically opaque words, such as *honeymoon*. Past studies have shown that this kind of incorrect use of morphological information and word-based association is one of the most frequent mistakes (Bensoussan & Laufer, 1984; Huckin & Bloch, 1993; Laufer, 1989; Nassaji, 2006). To manage this variability in the semantic transparency of word morphemes, it is important for readers to consider contextual meanings as well, because they provide a necessary semantic framework in a given scenario.

In this regard, previous studies have shown both the effectiveness and the difficulties of using both morphological and contextual information in inferencing (Brusnighan & Folk, 2012; Hamada, 2014; Mori, 2002; Mori & Nagy, 1999). In particular, such a combined use of linguistic clues is demanding when the learners' proficiency level is low (Hamada, 2014): learners often infer unknown words literally even when the surrounding contexts do not support their interpretations. The potential difficulties of inferring such opaque lexical items are related to more complex processes involved in successful inferences, i.e., noticing the semantic conflicts between morphemes and the context and then generating contextually appropriate meanings (Oakhill, Cain, & Nesi, 2016).

However, few studies have investigated the processes with which Japanese EFL readers have particular difficulties in inferencing because most previous studies have only examined the outcomes of inferences; as a result, little is known regarding the best means of supporting such learners. Therefore, the present study will examine the processes and interpretations of unknown words using morphological and contextual information, and will explore the necessary conditions or effects of instruction that enable their successful interpretations. To achieve these aims, we focused on intermediate- and beginner-level Japanese EFL learners, and addressed different issues that were unresolved for each proficiency level of the learners.

1.2 Organization of This Dissertation

This dissertation consists of the following six chapters. Introduction (Chapter 1), Review of Related Literature (Chapter 2), Study 1 (Chapter 3), Study 2 (Chapter 4), General Discussion (Chapter 5), and Conclusion (Chapter 6).

Chapter 2 reviews previous studies that are related to the current research. First, the roles of lexical inferencing during reading and its basic cognitive processes are introduced. Next, linguistic clues used for inferencing, especially morphological and contextual information, are described. Then, past studies that examined the combined use of the two information sources are explained. Further, processes involved in the use of the two information sources, the related factors affecting the process, and the methodologies used in the present study are explained. Finally, the summary and limitations of previous studies, as well as any unresolved issues unique to each of the intermediate- and beginner-level learners are discussed.

The current study includes a total of five experiments (Experiments 1–5). Study 1 included three experiments (Experiments 1–3) that examined intermediate-level learners' use of morphological and contextual information in the processing and interpretation of unknown words and contexts. Study 2 included two experiments (Experiments 4 and 5) to investigate beginner-level learners' use of morphological and contextual information in lexical inferencing and the effects of inferencing training were also examined. Figure 1.1 presents the overview of the five experiments.

In Chapter 3, Study 1 targeted intermediate-level Japanese EFL learners and examined the processing of unknown words with morphological and contextual information and its relationship with the interpretations of unknown words and the surrounding context. Experiment 1 explored whether Japanese EFL learners were sensitive to the semantic relationship between morphemes and context while reading as well as whether they could appropriately interpret unknown words according to the semantic transparency. In addition,

their semantic representation of the context sentence was examined. In the experiment, the participants read a single sentence context with either transparent or opaque unknown compound words (reading task), and then translated the context (translation task) into Japanese. The target reading times, interpretations of target words, and surrounding context in the translation protocols were analyzed.

Experiment 2 examined the relationship between text processing and semantic memory representations of unknown words. The participants read two-sentence passages with target words, and answered a comprehension question that evaluated their grasp of the meaning of the target words immediately after reading them. The reading times for the two sentences and their interpretations of the target words produced for the comprehension questions were analyzed.

Experiment 3 investigated their use of a lexical inference strategy and its relationship with the interpretations of unknown words. The participants were asked to infer the unknown words, while verbalizing their thoughts (think-aloud task). Their inferential strategy use and its relationship with their interpretations were examined. In addition, qualitative analysis was conducted to explore the causes of inappropriate interpretations for opaque targets.

In Chapter 4, Study 2 targeted beginner-level Japanese EFL learners to explore their use of morphological and contextual information and examine the effects of inference training. Experiment 4 explored whether they could understand the semantic relationship with morphological and contextual information and examined their ability to exploit context-related information to activate the semantic representations of upcoming words. The participants were asked to judge the semantic in/consistency between the two information sources, both with and without time constraints (on- and off-line consistency judgment tasks). The correct response rates and reaction times of the judgments were analyzed.

Experiment 5 explored the effects of inferencing training. The participants inferred target words using a step-by-step inference worksheet that described when and how to use

morphological and contextual information, and they also took pre- and post-lexical inference tests. The performances between the pre- and inference training and the pre- and post-tests were compared to examine the effects of training.

Chapter 5 discusses the findings from the five experiments based on learner proficiency. Chapter 6 includes the major findings of this study, its limitations, and suggestions for future research. Finally, this dissertation concludes with the pedagogical implications for lexical inferencing training and reading instruction aimed separately at intermediate- and beginner-level learners.

Study 1: Intermediate-Level Learners		**Measurements**
Experiment 1	Perception of semantic relationship between morphemes and context Interpretations of unknown words and context	Word-by-word self-paced reading task Translation task
Experiment 2	Semantic memory representations of unknown words and text processing in two-sentence passages	Sentence-by-sentence self-paced reading task Reading comprehension question
Experiment 3	Use of lexical inference strategy	Lexical inference task (Think-aloud method)
Study 2: Beginner-Level Learners		**Measurements**
Experiment 4	Understanding of semantic relationship between morphemes and context	Semantic consistency judgement task
Experiment 5	Effects of lexical inference training	Lexical inference tasks (Pre-test, inference training, post-test)

Figure 1.1. Overview of the five experiments in the present study.

Chapter 2

Review of Related Literature

2.1 Lexical Inferencing in Reading

2.1.1 Definition of lexical inferencing

Vocabulary knowledge plays a key role in successful reading comprehension (Horiba, 2012). A large vocabulary assists learners in having easy access to the content of a text. On the other hand, the presence of unknown words can influence accuracy and fluency of reading comprehension (e.g. Laufer & Ravenhorst-Kalovsk, 2010). This is more often the case with second or foreign language readers due to their limited vocabulary knowledge in comparison with their L1 counterparts.

Readers mainly employ three strategies in dealing with unknown words: ignore, consult (e.g. a dictionary, a teacher), and infer (Fraser, 1999; Paribakht & Wesche, 1999). Although many factors may influence the reader's decision on the strategy, such as keyness and salience of the word in text (Wesche & Paribakht, 2010), in cases where the reader finds it necessary to determine the word's meaning, their optimal choice is often to infer its meaning, i.e. lexical inferencing (Fraser, 1999; Paribakht & Wesche, 1999).

Lexical inferencing is defined as "making informed guesses as to the meaning of a word in light of all available linguistic cues in combinations with the learner's general knowledge of the world, her awareness of context and her relevant linguistic knowledge" (Haastrup, 1991, p. 40). This definition underscores the interaction of textual information and readers existing knowledge as "it is probably best to think of lexical inferencing as qualified guessing of the meaning of lexical items in context, rather than guessing from context, as contextual cues are only one of several knowledge sources" (Schmidt, 2010, p. 32). Thus, the inferencing process

shares much in common with reading comprehension in general: both require understanding textual information interactively with one's prior knowledge (Kintsch, 1994).

2.1.2 Merits and challenges of lexical inferencing

Success in lexical inferencing contributes to accurate reading by filling in a lexical gap in texts (de Bot, Paribakht, & Wesche, 1997), and can also lead to fluent reading (Wesche & Paribakht, 2010). In addition, although the primary purpose of inferencing is accurate understanding of a word and the text as a whole, the effortful process involved in inferencing would add to the reader's vocabulary knowledge. In an attempt to understand the unknown word, the reader may build word form and meaning associations, which is an essential first step toward "incidental" acquisition (Elgort, 2017; Huckin & Coady, 1999; Paribakht & Wesche, 1999; Swanborn & de Glopper, 1999).

This *incidental vocabulary learning* occurs as a by-product of reading, and it is compared with *intentional vocabulary learning*, where purposeful efforts are directed mainly to learn vocabulary (Nation, 2013). Huckin and Coady (1999) list the strengths of this learning pattern. First, it allows gaining a richer sense of the word from context that cannot be obtained from a typical word list. Second, it engages learners in reading and vocabulary acquisition at the same time. Third, it encourages individualized learning since readers can choose their own reading material. The merit of incidental vocabulary learning is generally recognized, as studies have shown that core meanings, collocational knowledge, and word usages are better learned from context through repeated and varied contextual exposures (Bolger, Balass, Landen, & Perfetti, 2008; Waring & Takaki, 2003; Webb, 2007). Especially in EFL contexts, where learners hardly receive linguistic input outside the classroom, it is recommended to consolidate vocabular knowledge in context, after initially learning it intentionally (Kadota & Ikemura, 2006).

Despite these potential advantages, chances of successful inferencing are found to be

surprisingly small, especially in L2 reading (Bensoussan & Laufer, 1984; Folse, 2004; Frantzen, 2003). For example, Nassaji (2003) reported that their participants made successful inferences in only 25.6% cases. Although this is sometimes because of lack of sufficient clues in a text (Beck, McKeown, & McCaslin, 1983), even when useful clues are available in the text, L2 readers sometimes fail to notice them, or misuse them (Laufer, 1989). In general, accurate inferencing is thought to be based on knowing approximately 95% or above of the running words of a text (Hu & Nation, 2000). This threshold level requires a minimal vocabulary size of 3,000 words or more (Huckin & Coady, 1999). Understandably, less proficient L2 readers tend to have very limited success in lexical inferencing. In addition, unsuccessful inferences may lead to incorrect understanding of the word, or even result in misinterpretations of other parts of text where the wrongly inferred meaning makes sense (Bensoussan & Laufer, 1984).

One way to avoid such wrong inferences is by using a dictionary. However, this may not be a perfect solution for several reasons. First, readers need to choose from the multiple senses in a dictionary entry to decide a word's meaning in the particular context (Verspoor & Lowie, 2003). Second, regularly consulting the dictionary may interfere with the reading process. Moreover, since readers only look up "new" words, they may dismiss "familiar" ones that they have mistaken for known words, even when they do not fit the surrounding context. This misinterpretation often occurs when the unknown words have similar forms to certain known words or when the words contain known morphemes that turn out not to be indicative of the words' meanings (Laufer, 1989).

Given the merits of lexical inferencing in reading comprehension and vocabulary development, it is of great pedagogical value to improve learners' ability to identify unknown words and to make successful lexical inferences. To reveal the mechanism of lexical inferencing and challenges for L2 readers in employing this strategy, the following section reviews the basic cognitive processes involved in lexical inferencing.

2.1.3 Cognitive processes in lexical inferencing

Drawing on Levelt's (1989) speech production model, de Bot et al. (1997) explained learners' meaning search process in lexical inferencing. This model places an emphasis on the role of lexical knowledge in both receptive and productive language use. In a lexicon, *lemma*, or a word's semantic and syntactic information, is mediated with *lexeme* (phonological and morphological information) and concepts. In receptive activities (i.e. reading and listening), processing lexeme activates related lemma, and then concept. In the case of unknown words, encountering an unknown word form (lexeme) creates an empty lemma. Readers can use both bottom-up (e.g. morphemes) and top-down (e.g., context, prior knowledge) strategies to fill in the missing lemma.

Huckin and Bloch (1993) proposed a tentative cognitive processing model of lexical inferencing. This framework consists of two components: generator/evaluator and metalinguistic control (see Figure 2.1). The generator/evaluator component generates and verifies hypotheses on the meaning of unknown words, drawing on various interconnected knowledge-based modules, including vocabulary knowledge, text representation, and prior knowledge. The metalinguistic control component controls the conscious decision-making process, such as generating and evaluating inferences. This decision-making process follows both serial and parallel patterns, and governs the entire inferencing process.

When the evaluator provides a positive evaluation (PE) for the inferred word meaning, the semantic and syntactic information is applied to update the reader's mental representation of the text, which also adds to his/her vocabulary knowledge. In contrast, when the evaluator gives a negative evaluation (NE), readers continue to generate and check another hypothesis when they are still motivated to infer the meanings of certain words.

This model emphasizes the critical role of readers' lexical knowledge and comprehension of the text being read, which can provide clues for inferencing, and metacognitive processing,

which may affect the way readers use knowledge sources in combination with inferencing strategies.

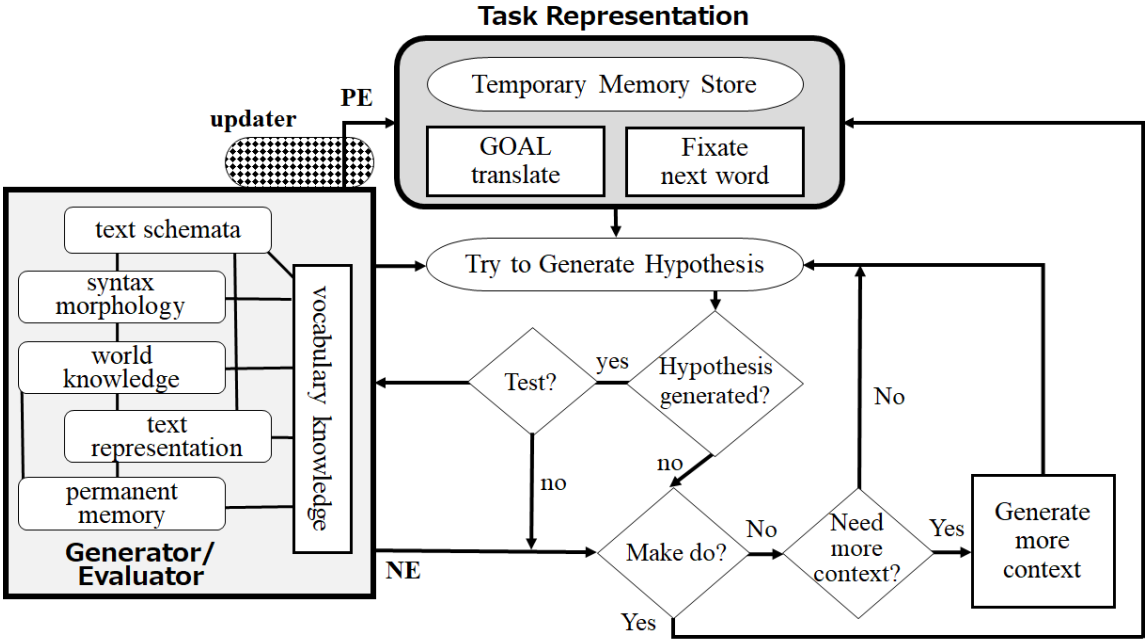


Figure 2.1. Cognitive model in lexical inferencing (adapted from Huckin & Bloch, 1993, p. 170). PE = positive evaluation, NE = negative evaluation.

However, in the real world, it should be noted that these inferencing and evaluative processes do not function perfectly: readers use certain sources of information more often than others and do not always sufficiently evaluate their inferences because they tend to make the minimum effort necessary to infer the meanings of unknown words (Huckin & Bloch, 1993; Mondria & Wit-de Boer, 1991). Therefore, once they think they could understand the meaning of the word, they are less likely to additionally evaluate it (Mondria & Wit-de Boer, 1991) or use the same information source for further inferences once they have used them previously (Huckin & Bloch, 1993).

2.2 Roles of Linguistic Clues Used for Lexical Inferencing

2.2.1 Taxonomy of knowledge sources

Previous studies have tried to capture the source of information that L2 readers use for lexical inferencing, known as *knowledge sources*, because generation and evaluation was made against them, and this was done by examining their verbal reports (Bengeleil & Paribakht, 2004; Chern, 1993; de Bot et al., 1997; Haastrup, 1991; Haynes, 1993; Huckin & Bloch, 1993; Nassaji, 2003, 2006; Paribakht & Wesche, 1999). These knowledge sources include morphological, grammatical, discourse, and L1 knowledge (e.g., Haastrup, 19991; Nassaji, 2003). These studies have shown that lexical inferencing draws from various linguistic (both the target language and their L1) and non-linguistic knowledge sources.

In the current study, we adopt the taxonomy offered in Wesche and Paribakht (2010). In this taxonomy, knowledge sources are broadly categorized into linguistic and non-linguistic sources (e.g. world knowledge), and the former is further divided into L1- and L2-based sources. However, the current study focuses only on L2-based knowledge sources in a text because it is often the lack of quality of textual clues that causes learners to fail to make appropriate inferences (Beck et al., 1983; Wesche & Paribakht, 2010). The definition of each knowledge source category is presented in Table 2.1.

In their taxonomy, further classification is made based on the language unit to which cues belong in a given text. These include word knowledge, sentence knowledge, and discourse knowledge. *Word knowledge* refers to cues within an unknown word itself (e.g. word morphology, word form); *sentence knowledge* refers to cues within a sentence containing unknown words (e.g. sentence meaning, sentence grammar); *discourse knowledge* entails cues beyond the sentence in which the unknown word is embedded (e.g. discourse meaning, formal schemata).

Table 2.1

Taxonomy of L2-Based Linguistic Knowledge Sources (Adapted From Wesche & Paribakht, 2010, p. 77)

Category	Knowledge Source	Definition
Word Knowledge	Word association	Association of the target word with another familiar word or network of words
	Word collocation	Knowledge of words that frequently occur with the target word
	Word morphology	Morphological analysis of the target word based on knowledge of grammatical inflections, stem, and affix
	Word form (written)	Knowledge of formal (orthographic or phonetic) similarity between target word, or a part of it
Sentence Knowledge	Sentence meaning	The meaning of part or all of the sentence containing the target word.
	Sentence grammar	Knowledge of the syntactic properties of the target word, its speech part and word order constraint
	Punctuation	Knowledge of rules of punctuation and its significance
Discourse Knowledge	Discourse meaning	The perceived general meaning of the text and sentences surrounding the target word
	Formal schemata	Knowledge of the macro structure of the text, text types and discourse patterns and organization

With regard to readers' actual choices, word knowledge (word morphology in particular) and sentence knowledge (meaning of the sentence) are favored over other clues. This is because they are easily accessible, possibly due to their locality (Haynes, 1993; Wesche & Paribakht, 2010), and, compared with rhetorical cues, word- and sentence-based cues offer more meaning-oriented clues for inferring the meaning of unknown words (Wesche & Paribakht, 2010). Since readers' inferencing process often ceases with perception, further search is not always necessary (Huckin & Bloch, 1993; Mondria & Wit-de Boer, 1991), if local context cues offer sufficient information. Only when readers fail to make successful inferences from immediate context cues

would they seek discourse-level information (Bengeleil & Paribakht, 2004). Furthermore, as readers tend to draw on cues active in their working memory, distantly located information would be less likely to be accessed (Pulido, 2003, 2009; Wesche & Paribakht, 2010).

These linguistic features and readers preference of cue usage would suggest that lexical inferencing performance is largely determined by the property of word and sentence knowledge in the text. Thus, the next section reviews the role and readers' actual use, focusing on morphological and contextual cues in lexical inferencing.

2.2.2 Morphological information

A morpheme is the smallest unit of a word in morphological analysis, and serves as an affix, a stem, or a root (Nation, 2013). *Free morphemes* can stand alone as words (e.g., *natural* in *unnatural*), while *bound morphemes* can be used only with other morphemes (e.g., *un-*). The two major varieties of word formation in English are derivation and compounding (Bauer, Liebar, & Plag, 2013). Derivation includes adding derivational affixes (i.e., suffix, prefix) to the lexeme, while compounding features combination of two or more free morphemes. Some morphemes carry semantic information of the word, which is inherent to the word's meaning (Nagy & Anderson, 1984). Thus, recognition of these known morphemes in unfamiliar words would assist in inferring the unknown words (Mochizuki & Aizawa, 2000; Nation, 2013).

Nagy and Anderson (1984) investigated how well an L1 English child could determine the meanings of unknown derived words (e.g. *misrepresent*) from their familiar base words, or "immediate ancestor" in their study (e.g. *represent*), by rating the degrees of semantic relatedness of target-based pairs. The target items included suffixed words (e.g. *frustration*/*frustrate*), prefixed words (e.g. *unknown*/*known*), compound words (e.g. *farmhand*/*farm, hand*), and idiosyncratic words (e.g. *prophesy*/*prophecy*). The 182,100 derived target words were categorized into six types (SEM 0–5) based on the degree of semantic

relatedness (see Table 2.2). Further division was made between SEM 0–2 (inferable with little or some contextual aid) and SEM 3–5 (hardly inferable from the base word).

Table 2.2

Categories and Definitions of Semantic Relatedness (Adapted From Nagy & Anderson, 1984)

	Definition	Examples
SEM 0	The semantic relationship between the target word and immediate ancestor is semantically transparent.	cleverness/clever
SEM 1	The meaning of the target item can be inferred from the meaning of its immediate ancestor with minimal help from context.	misrepresent/represent
SEM 2	The meaning of the target item can be inferred from the meaning of its immediate ancestor with reasonable help from the context.	everyday/day
SEM 3	The meaning of the target item includes semantic features that are not inferable from the meaning of the immediate ancestor without substantial help from the context.	collar bone/collar
SEM 4	The meaning of the target word is related to the meaning of its immediate ancestor, but only distantly.	colleague/league
SEM 5	There is no discernible semantic connection: the meaning of the immediate ancestor of no use in learning or remembering the meaning of the target word.	peppermint/pepper

Note. The word on the left is the target word; the one on the right is its immediate ancestor.

The results showed that the meanings of 139,020 words (76.3%) were within SEM 0–2, suggesting that word meanings are largely inferable from their constituent morphemes with some contextual support. Therefore, readers have a good chance in making correct inferences of the unknown words with proper knowledge of the words' morphemes.

Previous studies have provided empirical evidence for the effectiveness of morphological analysis in lexical inferencing. In Nassaji's (2003) study of inferencing strategies, use of morphological knowledge led to the highest success rate (35.7%) in inferencing among all

knowledge sources, including world knowledge (29.2%). For this reason, Frantzen (2003) argued against the use of nonsense words as target words in lexical inferencing studies. Although the use of nonsense and pseudo-words excludes the possible influence of prior knowledge, it rules out the use of morphemes (e.g. word stem) in the course of inferencing, leading to less ecological validity.

Using morphological information can also aid word learning. Contextual word learning occurs when readers' attention is directed to the association between formal and semantic features of the word (Huckin & Coady, 1999; Paribakht & Wesche, 1999). Thus, analyzing the morphological structure of a word with reference to its meaning is conducive to building the initial form-meaning relationship. Previous studies have shown that retention of vocabulary knowledge is superior when learners process formal properties of the word, such as word morphemes (Fraser, 1999; Hu & Nassaji, 2012). The benefit of using morphological information in vocabulary learning is further evidenced by findings in intentional vocabulary learning research (Wei, 2014), where the word-part method proved to be an effective strategy. Linking the meaning of a known morpheme to that of an unknown word elicits deeper and elaborative processing of form-meaning association (Craik & Lockhart, 1972).

However, there are instances where word-form analysis results in wrong inferences, especially when the meaning inferred from word form is unrelated to the actual meaning of words (e.g. Clarke & Nation, 1980; Frantzen, 2003; Huckin & Bloch, 1993; Nassaji, 2003, 2006). Laufer (1989) broadly referred to these problematic lexical items as *deceptively transparent words*, which she defined as "words which seemed to provide clues to their meaning but in fact did not" (p. 11). For example, morphological decomposition of the word *infallible* into morphemes (i.e. *in-fall-ible*) would likely to result in misinterpretation of the word. She also identified five types of deceptively transparent words: (a) words with a deceptive morphological structure (e.g. *outline*), (b) idioms (e.g. *sit on the fence*), (c) false friends, i.e.

formal resemblance between L2 and L1 word (e.g. *sympathetic* was mistakenly interpreted as *simpati* ["nice"] in Hebrew), (d) words with multiple meanings (e.g. *abstract*), and (e) synforms, i.e. pairs of words with similar form (e.g., *cute, acute*).

Regardless of the diverse forms of possible deceptions, misidentification is fundamentally caused by the reader/listener's false assumption that s/he knew the words but in fact did not. In fact, Laufer (1989) demonstrated that the participants were more likely to mistake unknown words for known words with deceptively transparent words than non-deceptively transparent words. Furthermore, misinterpreted words could serve as potential clues for other unknown words, which may lead to global misinterpretations (Bensoussan & Laufer, 1984). In addition, Nakagawa (2006) reported cases in which readers' initial correct lexical inference from context cues subsequently resulted in wrong inference by using morphological information.

These misidentification cases can be explained by the automatic nature of morphological decomposition (Pollatsek, Bertram, & Hyönä, 2011; Pollatsek, Slattery, & Juhasz, 2008). Visual word recognition studies have examined whether polymorphemic words (e.g. compound words, words with prefixes) are accessed via a *decomposition route* (i.e. retrieving lexical information via morphological composition of the word) or a *whole word route* (i.e. in which lexical information is directly retrieved from lexicon). Generally, a decomposition route is used at lexical level even for opaque words, the meaning of which cannot be derived from morpheme-based meanings (for a review see Hyönä, 2015). For novel words, where an established *whole word route* is lacking, readers may resort to a decomposition route to access the word meaning (Pollatsek et al., 2008, 2011). Thus, ready access to morphological information may give readers the false perception of knowing the words, this kind of automatic decomposition was also observed in lexical inference study (Fraser, 1999.

2.2.3 Contextual information

Contextual information is another potent source for inferring meanings of unknown words. Its major role in lexical inferencing is to put constraints on the possible interpretations. Given the polysemy of most words, it is context that helps determine the meaning of the word in the given context (Duffy, Morris, & Rayner, 1988; Elston-Güttler & Friederici, 2005).

However, an important caveat in the use of contextual information is variability in informativeness of the context (Beck et al., 1983; Webb, 2008). Some contexts can provide direct cues to pin down the meaning of a word, while others give little clue or offer only abstract meanings. According to Beck et al. (1983), contexts can be classified into four types: directive, nondirective, general, and misdirective (see Table 2.3). The proportions of correctly identified the meanings of words (replaced with banks) by their adult participants for each category were 3%, 27%, 49%, and 83%, respectively.

Table 2.3

Classification of Contexts (Adapted From Beck et al., 1983, pp. 178–179)

Category	Definition
misdirective context	contexts that seem to direct the reader to an incorrect meaning for a target word
nondirective context	contexts that seem to be of no assistance in directing the reader toward any particular meaning for a word
general context	contexts that seem to provide enough information for the reader to place the word in a general category
directive context	contexts that seem likely to lead the reader to a specific, correct meaning for a word

Unfortunately, most authentic contexts are not effective for cueing the meanings of unknown words (Beck et al., 1983; Bensoussan & Laufer, 1984; Frantzen, 2003). Bensoussan and Laufer (1984) investigated the guessability of 70 target words in a text and found that 29

words had no clues from context, while 13 words had clear contextual clues. In addition, Frantzen (2003) reported cases in which learners accurately inferred the meanings of unknown words presented in isolation, but interpreted the same words incorrectly in contexts. Although some of the inferred meanings in context warranted a logical interpretation, the result indicates that relying solely on context cues might be problematic in some cases.

With regard to vocabulary learning, however, contextual informativeness functions in an opposite way from its comprehension: that is, supportive contexts are ineffective for word retention, while contexts with moderate difficulty in inferencing are desirable for word learning. Supportive contexts enable readers to derive or infer lexical meanings easily from the surrounding text, thus they are less likely to attend to formal properties of the unknown words, resulting in less retention (Hu & Nassaji, 2012; Mondria & Wit-de Boer, 1991). A previous study showed that readers made less efforts to evaluate their inferences (Mondria & Wit-de Boer, 1991) and processed the formal properties shallowly (Hu & Nassaji, 2012) for words readily inferable from the context.

A more recent study examining the role of context reading for learning of new vocabulary also indicated that uninformative contexts were more beneficial for retention than were informative contexts (van den Broek, Takashima, Segers, & Verhoeven, 2018). It was suggested that the difficulty readers experienced in deriving word meaning from vague contextual cues forced them to retrieve semantic information from lexicon, while supportive contexts facilitated word comprehension, directing little attention to the word form. Given the incremental nature of vocabulary learning (Bolger et al., 2008; Waring & Takaki, 2003; Webb, 2007), it is important for readers to pick up available semantic information of a word derived from context in each encounter, in terms of both comprehension and learning.

2.2.4 Combined use of morphological and contextual information, and semantic

transparency of compounds

As reviewed, morphological and contextual information contributes to lexical inferencing in different ways. The former cues the semantics of a specific word while the latter also provides information inherent to a given scenario. However, readers are encouraged to draw from both sources to avoid the possible negative influence from semantically opaque and deceptively transparent morphemes and to complement ineffective contextual support. Rather than relying solely on one source of information, readers need to make judgment as to which and how much information to include in their interpretation of an unknown word based on its morphological properties and situated context.

In this regard, previous studies have investigated learners' combined use of morphological and contextual cues in lexical inferencing (Brusnighan & Folk, 2012; Hamada, 2014; Mori, 2002; Mori & Nagy, 1999). These studies have used compounds as their target words mainly because compounds are structured following relatively loose rules and vary in semantic transparency (Libben & Jarema, 2006). *Semantic transparency* is the degree to which the meanings of parts of a word are related to its whole meaning (Libben & Jarema, 2006), which lies on a continuum, ranging from fully opaque to fully transparent (e.g. Schäfer, 2018). However, compound words are generally categorized into the following three types for convenience: transparent compounds, semi-transparent compounds, and opaque compounds. Specifically, in *transparent* compounds, both constituents contribute to the word's meaning. For example, the meaning of *sunlight* can be inferred by combing the meanings of its constituent morphemes. In comparison, only one of the two constituent morphemes in *semi-transparent* words (or *semi-opaque* words) is related to the semantics of the word (e.g. *strawberry*). Finally, words like *honeymoon* are considered *opaque* compounds because neither of the constituent morphemes is indicative of the word's meaning.

Accordingly, while transparent words can be readily inferable from morphological

decomposition, semi-opaque and opaque compounds require readers to consider both morpheme- and context-based meanings. For this linguistic characteristic, the present study also adopts compound words as target words, and hence these studies were reviewed in some detail (see also Table 2.4).

Mori and Nagy (1999) investigated English-speaking students' interpretation of novel semi-transparent Japanese *kanji* compounds (e.g. 月食 [lunar eclipse]), and compared their performance under three conditions (i.e. *kanji*-only condition, context-only condition, *kanji*-plus-context condition). The participants were asked to choose the meanings of the target words from four choices (i.e. integrated [e.g. lunar eclipse], *kanji* distractor [e.g. a monthly meal ticket], context distractor [e.g. fireworks], anomalous [e.g. a traffic light]). The results showed that accurate interpretation (i.e. integrated) was highest when both *kanji* and context were presented, indicating the students' ability to draw from both sources in making proper inferences. Proficiency scores were correlated with the use of context and integration option. It was also indicated that metalinguistic awareness of morphological cues might influence the student's use of the information.

Mori (2002) examined the use of morphological and context cues, and its relationship with learner's belief in the use of these clues. In the experiment, English-speaking learners of Japanese (intermediate or pre-advanced level) were assigned the task of inferring the meaning of unknown words under three conditions, as in the above study. Unlike Mori and Nagy (1999), they had to produce inferred meanings by themselves using an open-ended format. Although the best performance was found in the kanji-plus-context condition, many participants over-relied on either kanji or contextual clues. The study also found that the information source students used for making inferences was related to their belief in the efficacy of the source (e.g. morpheme, context).

In the above studies, however, combined use of morphological and contextual information

(i.e. integrated) was always the most plausible inference due to the use of only semi-transparent targets; thus, the experiment design did not tap into the s*elective* use of morphemes. As semantically opaque words include unreliable morphological information, there are cases in which readers should completely discard morpheme-based interpretations. In this regard, the two studies below used both transparent (morphology reliable) and opaque (morphology unreliable) pseudo compound targets, whose semantic transparencies were manipulated by semantic in/consistency with surrounding context, which is identical with the current experimental design.

Brusnighan and Folk (2012, Experiment 2) investigated skilled L1 readers' processing of transparent and opaque words. Target pseudo compounds (e.g. *drinkblend*, *deskdoor*) were embedded in sentences in which literal interpretation of the target word was congruent (transparent conditions: *The party host used a blender to mix each guest a drinkblend last night.*) or inconsistent (opaque conditions: e.g. *The party host used a blender to mix each guest a deskdoor last night.*) with the context. After the reading task, the participants took a vocabulary test, in which they were asked to choose the meaning of the target word from two options: a context-based correct choice (e.g. a mixed beverage) and an incorrect choice whose meaning was unrelated to the correct choice (e.g. a cylinder of gas). The results showed that they made accurate responses for both transparent (94%) and opaque (89%) words in most cases, though more accurate responses were made with transparent target words. In addition, the longer reading time for opaque conditions was evidence for combining information from both sources during reading.

As for L2 learners, Hamada (2014) examined ESL learners' ability to use morphological context information for inferencing. In this study, leaners of different proficiency levels (beginning, intermediate, high-intermediate, and advanced levels), were subject to two conditions (morphologically reliable and unreliable conditions) in which they were asked to

infer the meaning of pseudo compounds consisting of an existing word and a pseudo-word (e.g. *rainfime*). In morphology reliable conditions, the meaning of one constituent of the morphemes (e.g. *rain*) was congruent with the context (e.g. *No one had an umbrella. We stood under the rainfime.*), while in morphology unreliable conditions, the meaning of the compound constituent was incongruent with the context (e.g. *The student is starting college this semester. She is buying a rainfime.*). The students were asked to infer the meaning of the target word by choosing from options that included the morpheme-based meaning, context-based meaning, two distractors, and *I don't know*. The results showed that beginner-level learners were more likely to make mistakes in choosing morpheme-based options in morphologically unreliable conditions, where morpheme-based interpretations were semantically inconsistent with the surrounding context. The over-reliance on morphological information of beginner-level learners was attributed to their lack of evaluation of the information from the two sources, neglecting context meanings.

Overall the results of inferencing studies on compound words highlight the importance of using both morphological and contextual information for making accurate inferences, and also the difficulty of tasks that require inferring without this information (Mori, 2002), as well as how the effectiveness of use of contextual information is related to learner proficiency (Hamada, 2014; Mori & Nagy, 1999).

Table 2.4

A Summary of Previous Studies on the Use of Morphological and Contextual Information in Lexical Inferencing

Study	Participants	Target words	Context	Condition	Answer type and scoring	Main findings
Brusnighan and Folk (2012, Experiment 2)	52 university-level native speakers of English	20 pseudo English compound nouns (e.g., *deskdoor*)	Informative	(a) transparent (b) opaque	Multiple-choice: (a) context-based, (b) unrelated (administered as a vocabulary test)	・Correct choice: transparent > opaque
Hamada (2014)	107 college-level ESL students (divided into beginning, intermediate, high-intermediate, and advanced groups)	20 pseudo English compound nouns (e.g., *rainfime*)	Informative	(a) morphology reliable, (b) morphology unreliable	Multiple-choice: (a) morphology-based, (b) context-based, (c) distracter, (d) distracter, (e) I don't know options	・Choice of morphology-based option in the morphology unreliable condition: beginning > high-intermediate, advanced learners (no difference between intermediate and the other three groups)
Mori (2002)	74 English-speaking Japanese learners (intermediate or pre-advanced level)	45 Existing semitransparent Japanese compound nouns adapted from Mori and Nagy (1999)	Neutral	(a) *kanji*-plus context, (b) *kanji*-only, (c) context-only	Open-ended: Scored on a 5-point scale	・Inference scores: *kanji*-plus-context > *kanji*-only, context-only conditions ・Integration of the two sources are affected by perceived efficacy of the strategy
Mori and Nagy (1999)	59 English-speaking Japanese learners (intermediate or pre-advanced level)	74 Existing semitransparent Japanese compound nouns (e.g., 月食 [lunar eclipse])	Neutral	(a) *kanji*-plus -context, (b) *kanji*-only, (c) context-only	Multiple-choice: (a) integrated, (b) *kanji* distractor, (c) context distractor, (d) anomalous	・Choice of integrated answer: *kanji*-plus context > kanji-only, context-only conditions ・Correlation between use of context and integration, and proficiency ・Individual differences in preference for certain information sources

2.3 Cognitive Processes in Flexible Use of Morphological and Contextual Information

The previous section reviewed the important role of using both morphological and contextual information in lexical inferencing, and proficiency-related difficulties in the practice (e.g. Hamada, 2014). Despite consistent success in experiment conditions, inferencing in normal reading is more difficult, given the large proportion of misidentified deceptively transparent items (e.g. Bensoussan & Laufer, 1984; Huckin & Bloch, 1993; Laufer, 1989). It is not uncommon for readers to make false interpretations for unknown words that are semantically opaque or deceptively transparent.

The difficulty in interpreting opaque words may arise from inconsistency between the literal meaning derived from word-based information (e.g. morphemes) and the actual meaning. Therefore, successful interpretation of opaque words requires more complex and flexible use of inferencing strategies in comparison with transparent words. To elucidate the particular sources of difficulties, and to explore effective pedagogical intervention, this section focuses on the cognitive processes involved in successful interpretations of opaque words.

To better inform our study with compound words, we also draw on findings of studies on idiom, homonym, and reading research. Idiom studies in particular have provided valuable insights as idioms share important linguistic characteristics with compounds, and the processing mechanism of idioms is well documented in L1 studies. An idiom (e.g. *a piece of cake*) is "a figurative expression that usually can be interpreted literally but that takes a nonliteral meaning when used in a specific context" (Cain, Oakhill, & Lemmon, 2005, p. 66). Therefore, when an idiom is used in its unknown figurative sense, readers need to reject the literal interpretation to obtain its figurative meaning, as in the case of opaque words. In addition, similar to compound words, the degree of semantic congruency between literal and figurative meanings (i.e. transparency) differs across idioms. For example, the figurative meaning of the idiomatic expression *to get away with murder* can be derived from a literal interpretation while the

meaning of a less transparent idiom, *to be wet behind the ears* (to be young and inexperienced), defies a literal inference (Cain et al., 2005). To unveil the cognitive process involved in lexical inferencing, this study adopts Levorato and Caccriari's (1995, 1999) Global Elaboration Model (GEM), which considers readers' search for figurative interpretations of idioms as their attempts to achieve a global and coherent representation of context as a whole, going beyond their literal counterparts. It mainly emphasizes the reader's ability to monitor comprehension, and to cultivate contextual information to make context-based inferences.

2.3.1 Perceiving potential inconsistency between morphological and contextual information

Since morpheme-based interpretations of opaque words are found unfitting for the context, readers need to perceive or notice the inappropriateness of the literal interpretation within the context if the meanings are unfamiliar to them. In GEM, this noticing behavior results from monitoring of one's emerging comprehension, which directs the reader's attention to context, as well as target lexical items (Levorato & Cacciari, 1995, 1999). Therefore, younger children, whose processing patters are more piece by piece, tend to dismiss the semantic inconsistency because they give less attention to context (Oakhill, Cain, & Nesi, 2016).

Lexical inferencing studies have also shown the important role of such metacognitive monitoring (Hamada, 2014; Hu & Nassaji, 2014; Huckin & Bloch, 1993; Nassaji, 2003, 2006). It would allow them to identify the difficulties in their inferences and help them explore necessary information that will lead to successful inferences.

With regard to readers' comprehension monitoring, past reading studies have employed the inconsistency-detection paradigm (IDP; e.g. Albrecht & O'Brien, 1993). In this paradigm, readers read a text with target information (e.g. *Mary ordered a cheeseburger*.) that contradicts a prior description (*Mary had been a vegetarian*.), and detection of the inconsistency can be

viewed as their successful coherence monitoring. Previous L2 studies have shown that they were able to detect inconsistency when the target and prior description were adjacent to each other, while it was demanding when they were separated by single (Morishima, 2013) or several sentences (Ushiro, Nahatame, et al., 2016) because of their limited cognitive resources available for monitoring the intersentential relations.

2.3.2 Prioritizing contextual meaning

Although GEM presupposes that the readers' perception of the inconsistency between morphological and contextual information prompts their search for an alternative interpretation (Levorato & Cacciari, 1995, 1999), it might not always be the case, especially for L2 readers. In a homonym study by Ushiro et al. (2010), the participants were asked to translate sentences containing target homonyms used in their unfamiliar senses. Some participants revealed in their translation protocols that they interpreted the target homonyms using their contextually inappropriate yet primary known senses even though they accurately translated the surrounding context. This finding was taken as evidence for readers' inflexibility in sticking to known word meanings. Thus, detection of inconsistency does not necessarily lead to a search for more appropriate interpretations.

In this regard, GEM also highlights the importance of readers' metalinguistic knowledge. This involves the awareness that literal meanings of lexical items are not always their actual meanings. For L1 readers, this awareness develops as their exposure to figurative, non-literal expressions increases; they come to know the linguistic fact that literal interpretations do not always accord with the real meaning and try to explore alternative meanings (Cain, Towse, & Knight, 2009; Levorato & Caccari, 1999). However, possibly because of the lack of linguistic exposure to the target language, this awareness may differ widely in L2 readers, as their personal efficacy of information sources (e.g. morpheme) may influence their decision as to

which information they should use for interpretations (Mori, 2002; Mori & Nagy, 1999). Thus, prioritizing context-based meaning over morpheme-based interpretation appears to be critical for successful inferencing of opaque words for L2 learners.

2.3.3 Making appropriate inferences

Having noticed the inappropriateness of the literal interpretation in context and determined to search alternative interpretation, readers then need to generate reasonable fitting interpretations. However, unlike transparent words, whose meanings can be derived from their constituents, the meanings of opaque words cannot be inferred from morphological cues. Therefore, successful interpretation of opaque words would require accurate comprehension of the text to generate context-based inferences (Levorato & Cacciari, 1995). In doing so, although readers can achieve appropriate interpretations solely from contexts, they may attempt to integrate information from both sources (Hamada, 2014; Mori & Nagy, 1999), the same as readers would draw on meanings of idiom components, which they then test in context to achieve a figurative interpretation (Cain et al., 2005).

Regarding readers' attempts to achieve a coherent representation of a text, IDP studies have also provided insights into how readers resolve inconsistent information in texts. In Ushiro, Mori, et al.'s (2016) study, EFL learners read texts containing inconsistent information and subsequently performed a written recall task. The recall protocols revealed that some participants maintained the coherence of the story by modifying interpretations of contradicting information in their memory. A think-aloud study by Ushiro et al., (2018) also showed that readers sometimes elaborated on the contextual information and added their own explanations about why inconsistent events occurred, without changing the interpretations of original text information. These findings raised the possibility that readers might make morpheme-based inferences for opaque words by changing literal interpretation of context.

2.4 Influencing Factors for Flexible Use of Morphological and Contextual Information

The previous section reviewed cognitive processes necessary for flexible use of morphological and contextual information based on semantic transparency of the unknown word. This involves (a) perception of the potential semantic incongruence between morphological and contextual cues, and (b) generation of contextually appropriate meanings, while suppressing unfitting literal interpretations. This section briefly reviews the potential influencing factors in the process, focusing on the characteristics of the reading material (i.e. context quality), individual differences (learner proficiency), and task requirements (purposes of reading).

2.4.1 Contextual informativeness

As mentioned in the previous section, the majority of words can have their meanings pinned down in context. However, due to variability in context quality, some contexts aid inferencing far better than others (Beck et al., 1983; Webb, 2008). Context quality in the current study was found to affect both readers' perception of the inconsistency between morphological and contextual information, and their subsequent meaning generation process.

Brusnighan and Folk (2012, Experiment 1) examined L1 readers' online integration of morphological and contextual information in the reading process. In their experiment, skilled L1 readers were asked to read two-sentence texts containing transparent or opaque compound words in informative or neutral contexts. Longer reading time was reported for processing opaque words only in informative contexts, indicating that the semantic incongruence of opaque words in neutral context was overlooked by the readers.

Context may also help readers pin down the accurate meaning of a word. Studies on lexical disambiguation in context suggested that both L1 and L2 readers initially activate multiple known meanings, but then the contextually appropriate meaning is selected,

suppressing unnecessary information by both L1 and L2 readers (Duffy et al., 1988; Elston-Güttler & Friederici, 2005). This holds true even when context-based meanings are not in the readers' lexicon. Ushiro et al.'s (2013) qualitative analysis implied that Japanese EFL readers made fewer primary-sense-based errors for processing homonyms when target words were embedded in more supportive contexts. The richer contextual information would have helped them prioritize contextually appropriate inferences and suppress alternative less relevant meanings than less directive contexts.

2.4.2 L2 reading proficiency

Proficiency is considered the key factor in determining readers' success in lexical inferencing and the way they use linguistic clues. Generally, skilled readers have been found to make more successful inferences in virtue of their better understanding of text (Bengeleil & Paribakht, 2004; Chern, 1993; Haastrup, 1991; Wesche & Paribakht, 2010). The current study focuses on the role of reading proficiency in three aspects: (a) the use of contextual information, (b) the use of inferencing strategy, and (c) flexibility in interpretations.

Readers' ability to exploit contextual information differs on the basis of proficiency. Hamada (2013) examined Japanese EFL learners' lexical inferencing ability in strongly and weekly constraining contexts, using a semantic relatedness judgment task. The results showed that high proficiency readers could activate specific senses of the unknown words in strictly constrained contexts while deriving general senses of unknown words from loosely constrained contexts. This difference indicated that they could narrow down the possible meanings of the unknown words according to context quality. On the other hand, less proficient leaners could only activate general senses of words regardless of contextual constraints. The result indicated that less proficient readers were less likely to benefit from context informativess in lexical inferencing. Bengeleil and Paribakht (2004) compared the use of inferencing strategies between

intermediate and advanced learners. Analysis of think-aloud protocols showed virtually identical patterns of cue choices between the two groups, while advanced learners attempted more successful inferences, showing their more effective cue use. Compared with intermediate learners who tended to draw on more knowledge sources, advanced learners were more likely to benefit from immediate contextual cues to achieve an accurate understanding. Because of the inability to make successful inferences from local information, intermediate readers had to search for additional information.

Successful interpretation of opaque lexical items requires readers' monitoring of the semantic relationship between morphological and contextual information. In this regard, the efficiency of monitoring and/or evaluative processes can be influenced by learner proficiency. If readers consume most of their cognitive resources for lower-level processing, such as word recognition and syntactic parsing, little attention is directed to higher-level processing activities such as comprehension monitoring (Grabe, 2009). Thus, younger children or less skilled readers often fail to notice semantic inconsistencies because of their limited processing capacities (Oakhill et al., 2016). University beginner-level L2 learners were also found to rely on literal interpretations of unknown words that were semantically incongruent with the context, and their undesired performances were attributed to a lack of attention to both morphological and contextual meanings (Hamada, 2014). Likewise, Nassaji's (2006) study indicated that not only did skilled learners use more monitoring and evaluating strategies, but also their use of such metacognitive strategies was more related to inference success than less skilled counterparts, suggesting that skilled readers are able to use these strategies more effectively.

In a study on homonym processing by Ushiro et al. (2010), the success rates for inferencing unknown secondary senses of target words showed no difference between proficiency groups in both open-ended (about 50%) and multiple-choice (about 90%) tests. However, error analysis of the test results suggested that proficient readers were able to replace

the morpheme-based primary interpretation with a contextually appropriate reading with the aid of contextual cues, while less proficient readers tended to stick to the primary senses. This result highlights proficient readers' flexibility in changing interpretations according to the given context.

In sum, proficient readers' effective and accurate understanding of contextual information enables them to monitor the inferencing process, to make flexible changes to their interpretations, leading to more successful inferences.

2.4.3 Reading purposes

Despite individual differences in proficiency levels, even the same reader might process unknown words differently (e.g. ignore, infer) depending on the situation. Since reading is a goal-oriented activity, and readers read texts for various reasons, such as for study, leisure, or for specific information (Grabe, 2009), the standards for successful comprehension and text processing are subject to their reading purpose in a given situation (Horiba, 2013). General situations, however, hardly require readers to pin down the exact meaning of unknown words unless they are of particular relevance to readers' purposes (Wesche & Paribakht, 2010). In fact, readers often ignore unknown words (Fraser, 1999; Paribakht & Wesche, 1999). In addition, trying to understand every novel word can affect reading fluency; hence making strategic decisions on important words is an element of efficient reading (Nation, 2009; Nuttal, 2005).

On the other hand, participants in lexical inferencing tasks are strongly motivated to reveal the meanings of target words, as it is the goal of the task. This is consistent with the assumption of the Involvement Load Hypothesis, which posits that the process depth of a word relies on the relevance of its meaning in a given task (Hulstijn & Laufer, 2001; Laufer & Hulstijn, 2001).

2.5 Assessment of the Use of Morphological and Contextual Information

This research project was meant to examine the inferencing of transparent and opaque unknown words with the aid of morphological and contextual information (e.g. Brusnighan & Folk, 2012). Given the multiple activities it involves (e.g. perception of the semantic relationship between information from the two sources, generation of context-based inferences), it is necessary to examine the process of inferencing behavior, rather than focusing merely on the outcomes of inferences. To this end, this section reviews methodologies adopted in current research, focusing on their operation, measurement, and data analysis.

2.5.1 Semantic consistency judgment task

The first step towards proper use of morphological and contextual information is to gauge the semantic relationship between cues from the two sources. To this end, we revised the semantic inconsistency detection task used in Koda (2000), so as to "measure the ESL participants' ability to integrate morphological and contextual information" (p. 308). In the computer-based task, participants were presented with a contextual sentence with the final word deleted (e.g. *People who never tell the truth are*), and were asked to decide as accurately and as quickly as possible whether the target words (e.g., *dishonest*, *honest*) that appeared on the next screen were semantically inconsistent with the preceding context. The correct response rates and reaction time were analyzed to examine the accuracy and efficiency of integration of morphological and contextual information.

Koda's study (2000) used the term *inconsistency detection*, as the correct answers for the experimental sentences were always the inconsistent one. However, the task designed for the current study was termed *semantic consistency judgment task*, as we are interested in whether participants were able to interpret the semantic relationship, or in/consistency, between morphological and contextual information for both transparent and opaque words.

2.5.2 Reading time

The yes-no judgment in the semantic consistency judgment task may explicitly direct participants' attention to morphological and contextual information; hence it is not suitable for assessing readers' perception during normal reading. In this regard, some past studies collected reading-time data from self-paced reading tasks to assess the online integration of multiple sources of information or implicit knowledge (e.g. morphological and contextual information) during reading (Brusnighan & Folk, 2012; Jiang, 2004, 2007; Oakhill et al., 2016). The processing time for target words or sentences was then compared between instances with literal and non-literal lexical items (Brusnighan & Folk, 2012; Oakhill et al., 2016), and grammatical or ungrammatical items (Jiang, 2004, 2007).

The longer reading time for non-literal items in comparison to transparent items was interpreted as successful integration of information from the two sources (Brusnighan & Folk, 2012; Oakhill et al., 2016). This interpretation bases itself on the assumption that if readers were able to monitor the comprehension of the semantic relationship between the two sources, they would experience processing difficulties with opaque lexical items because of the potential semantic conflict between information from the two sources. In this regard, an idiom study by Oakhill et al. (2016) mentioned that the inflated reading times for figurative expressions relative to literal counterparts might sometimes reflect the reader's further mental efforts to generate contextual interpretations, in this case figurative meanings.

2.5.3 Lexical inferencing task

The lexical inferencing task, where participants are given explicit directions to infer the target words, has been widely employed in previous lexical inference studies. The test format generally comes in two types regarding the way participants provide the answer: the open-ended and the multiple-choice format. In the former task, participants are asked to produce the inferred

meaning of the target word by themselves, whereas in the latter they are given the options to choose the one they deem most appropriate. This section focuses on the potential effects of the test format and scoring criterion that will allow us to estimate participants' inferential skills.

The multiple-choice format was employed in a number of previous studies on the processing of compound words (Brusnighan & Folk, 2012, Experiment 2; Hamada, 2014; Mori & Nagy, 1999), homonyms (Ushiro et al., 2010), and idioms (Oakhill et al., 2016). In these studies, options representing literal (e.g. morpheme-based meanings of opaque compounds, known meanings of homonyms, literal meanings of idioms) and appropriate options (e.g. unknown meanings of homonyms, context-based meanings of compounds, figurative meanings of idioms) were prepared to identify the source of information they used for interpretation, as well as causes of misinterpretations.

However, the provided options may give readers a hint in making inferences that they would fail to generate from their own reading experience. Thus, the use of options has the potential to bias readers' interpretation in positive and negative ways (Cain, et al., 2009). The bias can be particularly obvious in challenging tasks, e.g. processing opaque words, in which prioritizing contextual information and suppressing distracting morphological information were found most difficult (e.g. Ushiro et al., 2010). Therefore, to reveal a truthful picture of readers' own interpretations of the target words, the current study adopts an open-ended format in task design.

Scoring of lexical inference performance on an open-ended test was often based on the semantic similarities between the actual meaning of the target word when target words were real words (Mori, 2002; Nakagawa, 2006; Ushiro et al., 2010). However, it is often difficult to arrive at accurate meanings of opaque words or homonyms because the word-based information does not provide any information about their actual meanings, leading to underestimation of their inferential ability. In this regard, Ushiro et al. (2010) sorted the participants' incorrect

answers into primary meaning, context-congruent meaning, and unrelated meaning, in addition to correct (actual meaning) or incorrect scoring, to examine whether their mistakes were caused by inferential skills (unrelated) or over-reliance on primary meanings

On the other hand, previous studies with pseudo-transparent/opaque targets, as in the current study, used the multiple-choice format (Brusnighan & Folk, 2012, Experiment 2; Hamada, 2014). However, both studies used context-based options that were considered correct. Therefore, the present study regarded contextually appropriate interpretations as the intended correct answer, while categorizing participants' answers on the basis of the information source to identify the information sources used for interpretations, and causes of misinterpretations, as in Ushiro et al. (2010).

2.5.4 Think-aloud method

The think-aloud method has been widely used to examine the kinds of linguistic clues and metacognitive behaviors involved in readers' lexical inferencing (e.g. Bengeleil & Paribakht, 2010; Fraser, 1999; Haastrup, 1991; Hu & Nassaji, 2012, 2014; Huckin & Bloch, 1993; Nassaji, 2003, 2006; Paribakht & Wesche, 1999). This method requires participants to verbalize their thoughts while performing the task. The audial data are then transcribed, and categorized by inferencing strategies. Although the report would mainly reflect conscious and verbalizable thoughts, the data would allow us to explore into the readers' mental process with regard to the use of linguistic clues and metacognitive strategies.

Data analysis sometimes involves the comparison of frequency of strategy usage by different groups of participants to identify individual differences involved in strategy use (Bengeleil & Paribakht, 2010; Haastrup, 1991; Hu & Nassaji, 2012, 2014; Nassaji, 2003, 2006). In some reading research, readers' strategy use has been compared according to types of task instruction (Horiba, 2013; Ushiro et al., 2018) or text characteristics (Ushiro et al., 2018). Thus,

this method has potential to elucidate how the same individuals process and interpret unknown words with different characteristics, such as transparency/opacity of the words. Also, qualitative analysis of inferencing protocols would allow us to explore more specific questions: for example, how and why certain strategies were preferred over others; why a certain participant arrived at the correct or incorrect inference.

2.6 Summary and Limitations of Previous Studies and Relevance to the Current Research

Lexical inferencing in reading plays an important role in comprehension and vocabulary development. In making inferences, readers draw on information from various knowledge sources available in the text, resorting mainly to morphological and contextual cues. However, over-reliance on morphological information may result in erroneous inferences, due to semantic opacity or deceptive transparency of certain words (Bensoussan & Laufer, 1984; Laufer, 1989). Therefore, readers need to consider both morphological and contextual cues in inferring unknown words (especially opaque words). Past studies with compound targets showed the difficulties and proficiency-related differences of combined use of the two sources in lexical inferencing, especially semantically (semi-)transparent items (Hamada, 2014; Mori, 2002; Mori & Nagy, 1999). However, some limitations remain as to the insufficient interpretations of the findings, and unaddressed topics in previous studies.

The current research aims to examine lexical inferencing by EFL learners at different proficiency levels, with a focus on revealing the processing patterns and difficulties unique to readers at each proficiency level. Based on the findings of Hamada (2014), this study assumes that the threshold level regarding effective use of the two sources in inferencing lies between intermediate- and beginner-level learners. Below are the limitations of previous studies that examined the use of morphological and contextual information in general, and also problems specific to intermediate- and beginner-level learners, respectively.

Limitations of studies on combined use of morphological and contextual information

First, it is apparent that semantically opaque words are difficult to infer and that they often lead to literal word-based interpretations, but little is known about what process EFL learners have particular difficulties with: that is, whether it is failure (a) to notice the semantic inconsistency between morpheme-based and context-based meanings, or (b) to generate context-based meanings. While the former is possibly due to their lack of attention to contextual meanings or evaluation of the two sources (Oakhill et al, 2016), the latter is caused by their persistence in interpretation (Ushiro et al., 2010) or inability to use context to make inferences. To this end, the present study assesses processes as well as outcomes of reader interpretation, which will inform educators' focus of instruction.

Second, most past studies have tapped into participants' interpretations of target words via a multiple-choice task (Brusnighan & Folk, 2012; Hamada, 2014; Mori & Nagy, 1999). However, it is possible that the options reduced their need to generate their own inferences, thus biasing the inference generation process (Cain, et al., 2009), resulting in overestimated performance. This is especially significant for words with unrelated morphological structures because, having perceived the semantic incongruity involved in the opaque word cases, the readers were less inclined to choose the literal interpretations in the presence of other plausible options. As a result, this task could fail to properly assess their meaning generation process. To address this limitation, the current study employed an open-ended task to more accurately tap into readers' attempted interpretations.

As for the context characteristics, the context quality can potentially affect both perception of semantic relation (Brusnighan & Folk, 2012), and subsequent meaning generation process, suppressing unnecessary morpheme-based interpretation (Ushiro et al., 2013). However, past L2 studies have used only one type of context (Hamada, 2014; Mori & Nagy, 1999). Thus, the present study explored the effects of context informativeness on the inference

processes, and also its relationship with learner proficiency, since the effective use of context can be mediated by learner proficiency (Bengeleil & Paribakht, 2004; Hamada, 2013).

Finally, less is known about the inferencing process that leads to successful or failed attempts in processing opaque words. Identifying what and how strategies are used in the course of inferencing is important since it is not the use of a certain strategy but when and how to use certain strategies in a given situation that leads to successful inferences (Hu & Nassaji, 2014; Nassaji, 2006). Previous studies have reported cases in which inferences resulted in incorrect word-based interpretations, even though they might have perceived the semantic inconsistency between word- and context-based meanings (Mori, 2002; Ushiro et al., 2010). However, focusing on inference answers alone failed to address the underlying cause for and cognitive processes involved in the observed referencing patterns. Therefore, data collected using a think-aloud method have potential to elucidate this process, with insights on the role of individual differences in perception patterns.

Limitations of studies on intermediate-level learners

First, although past studies have investigated interpretations of unknown words via lexical inference task, and suggested that immediate or more advanced learners are generally able to use morphological and contextual information (Hamada, 2014; Mori & Nagy, 1999), given the relatively shallower engagement in unknown words during normal reading (Fraser, 1999; Wesche & Paribakht, 1999), it is necessary to examine how reading for comprehension operates.

In addition, lexical inferencing studies used excessively assessed readers' interpretations of target words only. Therefore, the type of semantic representations of word that readers have constructed as a result of reading is unclear. In addition, past studies have reported cases in which wrong identification of an unknown word as known resulted in distortion of subsequent

context (Bensoussan & Laufer, 1984), in which readers changed the interpretations of context so that inconsistent textual information made sense (Ushiro et al., 2018; Ushiro, Mori, et al., 2016). These results suggest that the presence of opaque unknown words would affect the interpretation of context during reading. Examining these underexplored issues will inform us on how processing opaque and other deceptively transparent words might affect reading comprehension in general, not just word-level problem.

Limitations of studies on beginner-level learners

Unlike intermediate-level learners, beginner-level learners have been reported to make frequent incorrect inferences of opaque words based on morphemes even in a lexical inference task, possibly due to their less attention to contextual meaning (Hamada, 2014). Thus, leaving them to their own devices to generate inferences of unknown words would be too demanding.

Therefore, what is necessary for such learners is to devise intervention that support them make successful inferences, considering both morphological and contextual information. In this regard, some studies have examined the effects of lexical processing strategy in general (e.g. Fraser, 1999), and also proposed a general lexical inference procedure (Clarke & Nation, 1980). However, given the multiple, complex processes required for inferencing opaque items, an ideal intervention should have particular emphasis on the processes that beginner-level learners have difficulties with. Accordingly, investigation into beginner-level learners will be aimed at (a) identifying processes that are problematic for them, and (b) examining the effects of instruction based on a better understanding of their inferencing process.

Overview of the current research

To address the above-mentioned insufficiencies, a total of five experiments were conducted in the current research project. Study 1 focused on the intermediate-level learners,

and Study 2 on beginner-level learners.

As for the experimental materials, this study employed pseudo compounds, whose semantic transparency was either transparent or opaque by manipulating the semantic consistency with the surrounding context. To examine the context effects, two types of contexts were prepared: informative neutral contexts, which were somewhat compatible with the directive context, and the general context of Beck et al.'s (1983) context classification.

Although derived words, like prefixed words (e.g. *miscircle*), also have morphological structures, whose morpheme-based semantic (e.g., *mis-*, *circle*) are conducive to inferring the meanings of the words, the present study did not use them for the following reasons. First, compared to compounds words, prefixed words have relatively straightforward semantics and are rule-governed; for example, the word *miscircle* "means that someone did X wrongly in some way" (Hyönä, 2015, p. 126); thus, they have less variability in semantic transparency. Second, our interest lies in whether readers could consider both morpheme- and context-based meanings and interpret them according to the transparency/opacity. Therefore, it was necessary for them to have ready access to the morphological structures and their semantics. As for this, word recognition studies have shown that word length could affect readers' morphological decomposition (Hyönä, 2015), in which the morphological structure of derived words is considered less salient. Also, learners' familiarity with affixes is greatly affected by their vocabulary size and educational settings (Mochizuki & Aizawa, 2000). Thus, it was difficult to practically control such variables, especially for less-skilled L2 learners.

Among compounds, some studies have used semi-transparent words to examine their ability to make integrated interpretation of the two sources (Hamada, 2014; Mori, 2002; Mori & Nagy, 1999). However, the current research focused only on fully transparent and opaque meanings, following Brusnighan and Folk (2012, Experiment 1), who examined online integration of morphological and contextual cues by comparing the processing of transparent

and opaque targets. In addition, as in deceptively transparent words, there are many cases in which word-form-based interpretations are hardly related to the actual meaning of the words (Bensoussan & Laufer, 1984; Laufer, 1989), like opaque words. Since such misidentification for these words shapes one of the most frequent mistakes in inferences (Laufer, 1989), the present findings can be applied to other lexical items, such as homonyms, and words with formal similarity with their known words.

Study 1 involved three experiments (Experiments 1–3) that examined intermediate learners' processing and interpretation of unknown words based on morphological and contextual information, and their effects on text comprehension. Experiment 1 examined whether intermediate Japanese EFL learners were sensitive to the semantic relation between the two sources, and were able to interpret unknown words appropriately according to the semantic transparency of the target word, to identify the difficulties readers experienced in the process. To examine the former process, a word-by-word self-paced reading was employed because it "allows one to operationalize automaticity as a dichotomy" (Jiang, 2007, p. 12). In addition, their sentence representation was analyzed via a translation task to see how they would achieve a coherent representation for sentences containing opaque unknown words. Experiment 2 focused on the semantic memory representation of unknown words as a part of text comprehension, and its relation to on-line text processing. Unlike Experiment 1, the participants read the text sentence-by-sentence to reveal what processes (e.g. noticing the inconsistency, inferring contextual meanings) they were engaging during reading (Oakhill et al., 2016). Experiment 3 examined their performance in a lexical inferencing task, and investigated what kind of inferencing strategies or behaviors would lead to successful or unsuccessful inferences for opaque unknown words based on their think-aloud reports.

Study 2 consisted of two experiments (Experiments 4 and 5) to investigate beginning learners' lexical inferencing performance and effects of training practice. Experiment 4 was

conducted to see whether beginning learners could build the semantic relationship with morphological and contextual information, and to determine the process Experiment 5 focused on improving their inference performance and examined the effects of lexical inferencing training, focusing on the use of morphological and contextual information in inferencing.

The current research project attempted to examine the use of morphological and contextual information in the processing and interpretation of unknown words. Giving a full account of the critical role of proficiency, it targeted different referencing processes of intermediate- and beginner-level learners. However, an examination based on a series of cognitive processes (e.g. perception of the semantic relation, generation of contextual meanings) would allow us to compare findings across studies and gain a better understanding of the characteristics of L2 lexical inferencing.

Chapter 3

Study 1: Use of Morphological and Contextual Information in the Processing and Interpretation of Unknown Words and Their Effects on Text Comprehension by Intermediate-Level Japanese EFL Learners

3.1 Experiment 1: Perception of Morpheme-Context Relationships, and Interpretations of Unknown Words and Contexts

3.1.1 Purposes, overview, and research questions of Experiment 1

Chapter 2 reviewed the importance and difficulties of using both morphological and contextual information in inferencing unknown words, especially words with semantically opaque morphological structures. Generally, successful understanding of these words requires readers' perception of the semantic relation of these two sources of information, and subsequent generation of context-based meaning. In these processes, learner proficiency in the target language plays an important role: more advanced learners, those at approximately the intermediate level or higher, are able to take into consideration the two sources and go beyond the literal understanding of morphemes, while this task is more demanding for less skilled readers, such as beginner-level learners (Hamada, 2014; Mori & Nagy, 1999).

However, the above investigations with the multiple-choice lexical inference task might by their nature have overestimated their performance, especially in the meaning generation process (Mori, 2002; Ushiro et al., 2010). In addition, perceiving the semantic relationship between the two sources would be more challenging during normal reading because less attention is required for unknown words than in a lexical inferencing task. Therefore, it is important to reexamine whether intermediate-level EFL learners are sensitive to these semantic relations during reading and are able to interpret unknown words appropriately according to the semantic transparency of the morphemes.

Furthermore, past studies have mainly analyzed only the interpretations of unknown words. Therefore, what kinds of sentence representations readers constructed remains unclear. Due to semantic conflicts, successful inferencing of (semi) opaque words requires revising some part of the word (morpheme) information. In this regard, past studies reported cases in which readers' wrong morpheme-based interpretations made them distort subsequent text meanings (Bensoussan & Laufer, 1984). This raises the possibility that the revision of linguistic clues in the pursuit of coherence could also extend to the surrounding context. Therefore, it is necessary to investigate how readers try to build the coherent representation of a sentence, focusing on what types of lexical and textual information they revise or distort.

In the experiment, sentential contexts including pseudo-compound target words whose morphemes were either semantically consistent (transparent) or inconsistent (opaque) with their surrounding contexts were used. The contexts were either informative or neutral for target word meanings to explore the effects of contextual support.

To compare the on-line processing, the present experiment adopted a word-by-word self-paced reading task because it was suitable for examining a binary proposition as to whether EFL readers could integrate morphological and contextual information on encountering unknown words (cf. Jiang, 2004; 2007). The participants read the context sentences word by word, and the reading times for transparent and opaque words were compared (reading task). Afterwards, participants translated the contexts including target words into Japanese to examine what kind of information readers used for their interpretations (translation task), and also the types of information revised through examination of the entire context sentence. These points as addressed in this experiment are summarized in the following three research questions (RQs):

RQ1-1: Are intermediate-level Japanese EFL readers sensitive to semantic relationship between morphological and contextual information during reading?

RQ1-2: What types of information do intermediate-level Japanese EFL readers use to interpret transparent and opaque unknown words?

RQ1-3: What types of information do intermediate-level Japanese EFL readers revise to achieve coherent representations of a sentence with opaque unknown words?

As for the above two processes (i.e., perception of semantic consistency, generation of context-based meanings) involved in successful interpretations of opaque words, the latter meaning generation process would be more demanding due to the adoption of an open-ended format in this experiment.

Context quality was expected to affect both processes. A study of L1 skilled readers with a similar experimental design showed that an on-line reading time difference was observed in only informative contexts and not in neutral ones (Brusnighan & Folk, 2012). In the meaning generation process, a biasing context can help suppress unnecessary information and prioritize context-based meaning (Ushiro et al., 2013).

3.1.2 Method

3.1.2.1 Participants

The participants of Experiment 1 were 30 undergraduate and graduate students from various majors. They had learned EFL in Japan for more than six years. Their estimated English proficiency was approximately at intermediate levels based on self-reported scores on English proficiency tests, such as the TOEIC Listening and Reading test (n = 5, M = 706.00, SD = 54.93, range = 620–765), TOEFL ITP test (n = 3, M = 490.33, SD = 22.03, range = 469–513), and EIKEN test (3rd grade: n = 6; pre-2nd grade: n = 1; 2nd grade, n = 6; pre-1st grade: n = 2).

Moreover, to estimate and compare the participants' reading proficiency in the present study across experiments, a reading proficiency test was prepared. It consisted of the reading

subsection of an EIKEN test. This test was chosen because it has been widely used in Japan and can be used to tap into differing levels of reading proficiency by manipulating the grade levels. One passage (k = 4) was adapted from the pre-2nd grade, three passages (k = 15) from 2nd grade, and one passage (k = 5) from pre-1st grade (Obunsha, 2000, 2001a, 2008). The time allocated for this test was 20 minutes. The average score was 15.20 (SD = 2.64), ranging from 11 to 21 (Cronbach's α = .81).

3.1.2.2 Materials

Target words

Initial target words consisted of 40 pseudo compounds that included two noun-based free morphemes. While three of them (i.e., *designbook*, *floorcover*, *fruitsweet*) were adapted from Brusnighan and Folk (2010), the others were created by the author because the target words in their study were considered difficult for EFL learners due to the low frequency of the constituent morphemes. The target words were created by replacing one or two of the constituent morphemes of existing compounds with synonyms. The existing words were chosen with reference to a dictionary of English compounds (Shintomi, Yamane, & Watanabe, 1998). In order for the participants to easily have access to the semantic representation of the target words, both morphemes were high-frequency words (Level 1 or 2 in the *Japan Association of College English Teachers* [JACET] *8000 list* [Ishikawa et al., 2003]). For example, an existing compound word *price war* was replaced with the target word *pricebattle* by changing the second constituent. While some compound words are separated by a space or a hyphen, this study consistently deleted these.

A pilot study was conducted to ensure that the participants would not have difficulty understanding the literal meanings of the target words. The participants were nine undergraduates who did not participate in the main experiment. They were presented with the

40 target words in isolation and were asked to (a) write down one possible meaning of each target word in Japanese, and (b) evaluate how easily they could come up with the meanings on a five-point Likert scale (1 = *very easy*; 2 = *easy*; 3 = *neither*; 4: *difficult*; 5 = *very difficult*). Based on the results, eight target words were excluded because they elicited multiple different meanings from some participants or had meanings that were difficult to construe (mean ratings over 3 [neither]). Consequently, the remaining 32 words were used as the target words in the main experiment (*M* = 1.87, *SD* = 0.33, *Min* = 1.41, *Max* = 2.43). The complete list of target words and their creation process is given in Appendix 1. The mean number of letters and syllables of the target words were 9.34 (*SD* = 1.15, *Min* = 7, *Max* = 11) and 2.47 (*SD* = 0.67, *Min* = 2, *Max* = 4), respectively.

Contextual sentences

For each target word, two types of sentential contexts were developed that provided different amounts of contextual information before the target words appeared: one context provided more information (informative context), while the other contained less information (neutral context). Informative contexts provided specific information pertaining to the target words. However, there were still some possible candidates to identify which source (i.e., morphemes or contexts) readers used for interpretation. For example, in the informative context *The shop closed after the hard* _____, the blank can be filled with events or reasons that caused the shop's closure, such as recession or lack of staff. On the contrary, in neutral context contexts, only abstract or more general meanings of the upcoming words could be inferred. For instance, for the context *Peter studied the reasons for the* _____, any action or event could be possible. Since topic familiarity could affect inference success (Pulido, 2003), most story events concerned daily-life topics so that EFL learners at a university level were able to understand without using any technical knowledge. To rule out the wrap-up effect (whereby final word of

a sentence is read longer to integrate the sentence proposition; Jian, 2012), every context had at least three words after the target words' appearance. The contexts consisted of high-frequency words (Level 3 or below in the *JACET 8000*). The average sentence length was 12.21 words (*SD* = 2.37, *Min* = 8, *Max* = 16). Since pseudo words do not have actual meanings, this study defined the transparency as the semantic in/consistency between target words and their surrounding context (Brusnighan & Folk, 2012). Therefore, in the above contexts, the target words functioned like transparent words, as their morphological information was consistent with the contexts.

Next, to prepare the opaque targets, the following manipulation was done. The 32 target words were divided into 16 pairs of two target words. Care was taken that the meanings of one target word did not semantically fit the contexts created for the other word. For example, for the pair of target words *pricebattle* and *businessbag*, the two contexts designed for *pricebattle* did not allow for *businessbag* to appear, while *pricebattle* was not semantically possible in *businessbag*'s contexts. Table 3.1 provides examples of the paired target words and contexts.

Table 3.1

Example of a Pair of Target Words and Context in Experiment 1

	Informative context	Neutral context
Transparent	The shop closed after the hard **pricebattle** in the area.	Peter studied the reason for the **pricebattle** on the Internet.
	After arriving at the office, Tim realized he had left his **businessbag** on the train.	When William was cleaning the house, he found a **businessbag** but it wasn't his.
Opaque	After arriving at the office, Tim realized he had left his **pricebattle** on the train.	When William was cleaning the house, he found a **pricebattle** but it wasn't his.
	The shop closed after the hard **businessbag** in the area.	Peter studied the reason for the **businessbag** on the Internet.

Note. Target words are presented in boldface.

Additionally, this combination of target words served to more closely minimize the variability of the target words' characteristics (i.e., the number of letters, syllables). Pairing target words meant that each word had four contexts conditions, crossing 2 (transparency: transparent, opaque) × 2 (context: informative, neutral context). Four presentation sets, in which every target word appeared only once for each condition, were then created to counterbalance the target words' presentation conditions across participants. Thus, all the participants would read all 32 target words in one of the four conditions, and encounter eight target words in each of the four conditions. Indeed, a 2 (transparency) × 2 (context) analysis of variance (ANOVA) of both the number of letters and syllables showed that there were no significant differences among the four sets (all *p*s > .10). This ensures that any reading time differences observed in this experiment can be considered the result of the treatments (i.e., semantic transparency, context informativeness).

Norming study

The validity of transparency and the informativeness of the context in relation to the target words were confirmed and revised through a norming study (i.e., semantic fitness judgment, informativeness judgment). The participants of this norming study were two Japanese graduate students majoring in English language education.

In the semantic fitness judgment, the raters were asked to judge whether or not the morpheme-based meanings of target words semantically fit the surrounding contexts. This was carried out to confirm that transparent words would receive "Yes" (semantically plausible), versus "No" (semantically implausible) for opaque words. In the informativeness judgment, the raters judged whether the concrete semantics could be derived from informative contexts, but there were still several possible meanings. For the neutral contexts, the judgment was whether only abstract meaning or general category can be derived.

Most of the items received the intended responses from the two raters (Semantic fitness judgment: 93.7%; Informativeness judgment: 87.5%). Then, contexts whose judgments received unintended responses by at least one rater were revised by the author; the two raters then reviewed the revised context and their agreement was obtained.

In addition, in order to distract participants from the experimental purposes, 32 filler sentences that did not contain any target words were also constructed. Of these, 22 sentences were yes-no comprehension questions asking about the literal meanings of the context to verify whether the participants had read the contexts for comprehension. The list of experimental and filler sentences is listed in Appendix 1.

3.1.2.3 Procedure

The experiment was conducted individually in a silent room. First, the author explained the general purpose and overall procedure of the experiment. Participants were then randomly assigned to one of the four presentation sets.

The reading task was administered using SuperLab 5.0 and Response Pad RB-740. The participants read the contexts word-by-word at their own pace while their reading times were recorded. They were instructed to read to answer the comprehension questions as quickly and accurately as possible (Jiang, 2004, 2007). Each trial began with a "Ready?" screen. When participants were ready, they were asked to push the "Yes" button. The next screen showed the sentence starting position with an asterisk and the word length of each word using underlines. At this time, each context was presented on one line. When participants pushed the button again, the asterisk disappeared and the first word appeared. After this, every time participants pushed the button, the subsequent word appeared and the preceding word disappeared. This meant that there was always only one word on the screen (non-cumulative presentation). Participants were also not able to read the preceding word again. When they reached the sentence-final position

and pushed the button, they proceeded to the next trial. In the case of filler sentences with comprehension questions, the question appeared on the next screen; participants were asked to answer it by pressing either the "Yes" or "No" button. In order to motivate their careful reading, feedback was given (○ or ×) after they responded. This procedure was repeated until all 64 sentences (i.e., 32 experimental and 32 filler sentences) were read. Figure 3.1 shows the sequence of a single trial in the reading task. To familiarize participants with this procedure, they performed three practice trials before the main session.

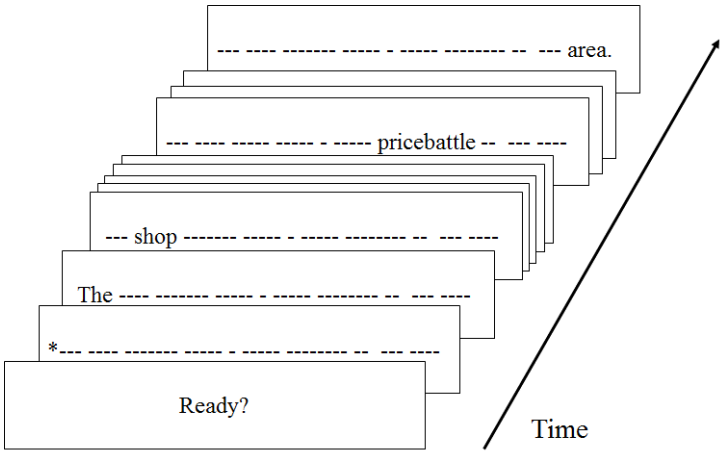


Figure 3.1. An example of the reading task in Experiment 1.

After a short break, the participants carried out the translation task. All 32 experimental sentences in the reading task were presented again in a booklet, but the presentation order was randomized across different presentation sets. Participants were instructed to translate the entire context including the target words into Japanese. After completing the task, the participants were informed that the target words were not real English words. The entire procedure lasted for approximately 70–90 minutes.

3.1.2.4 Scoring and analysis

As for the reading task, reading times for the target word were used for analysis. Before analysis, reading time outliers from the reading task were treated. Outliers were identified using both *SD*s from the participants' mean reading times and absolute values (Jiang, 2007). However, in this study, the reading times for target words were much longer than normal words in context, due to longer word lengths and novelty effects (i.e., novel words take longer to read than lexicalized words). Thus, unlike Jiang, the present study used each participant's mean reading times for target words instead of the mean of all words in context to calculate *SD*s. Reading times longer than 2*SD*s were substituted with the value, and those shorter than 200ms were regarded as indicating that the readers skipped the word and were discarded. In sum, these treatments accounted for 7.2% of the data.

To address RQ1-1, a 2 (transparency: transparent, opaque) × 2 (context: informative, neutral context) ANOVA was applied to the reading times of the target words. Again, if readers spent more time on opaque targets than on transparent targets, that would suggest that they felt processing difficulty in integrating both the morphological and contextual information while reading, showing their sensitivity to the semantic relationship between morpheme and context.

To answer RQ1-2, the translation protocols corresponding to target words in the translation task were categorized from two perspectives: (a) the information used for interpreting the target words and (b) whether the translation fit the overall contextual meaning. As a result, the following five translation types were ultimately identified (the following example protocols shown in parentheses were those produced for the target word *businessbag*). *Morphology-based interpretation* (MBI) refers to when participants simply translated the two semantic meanings of the two morphemes (e.g., 仕事用のカバン [business bag]). Although readers might have inferred the same semantic concepts only through the contexts (especially in informative contexts), this study regarded them as MBI. *Partially morphology-based*

interpretation (PMI) refers to interpretations where part of the morphemes' semantic information was deleted (e.g., ビジネス [business]), modified (e.g., 経営難 [financial difficulties]), or if some information was added (e.g., ビジネスバッグの売れ行き [sales of business bag]). This category was included because this interpretation was made based on part of the morphemes but remained consistent with contextual meanings; thus, this interpretation functions differently from the other categories. *Context-based interpretation* (CBI) refers to translations whose meaning was only related to the contexts and no morphemic information was included (e.g., 消費増税 [increase of consumption tax rate]). *Inappropriate interpretation* (INI) covers cases where the word meanings were semantically inappropriate for the context apart from MBIs. Thus, when participants used part of the morphemes but the interpretation was not semantically appropriate, the responses were categorized as INIs not PMIs. Finally, *None* refers to cases where the target words were not translated.

Regarding the relationship between the above interpretation types and their appropriateness, appropriate interpretation in this study was judged according to the semantic transparency of the target words. Since pseudo target words do not have pre-existent meanings, this study consistently judged the appropriateness of the interpretation based on whether the interpretation fit the context semantically (Brusnighan & Folk, 2012; Hamada, 2014). Accordingly, for transparent words, MBI, PMI, and CBI were considered appropriate, although MBI meanings were more concrete than the others. For opaque words, PMI and CBI were deemed contextually appropriate interpretations. PMI was included because it was assumed that readers might have noticed a discrepancy between the morphemes' meanings and the context, and thus made some revisions in order to generate a meaningful interpretation. The relationship between types of interpretations and interpretation appropriateness, along with definitions of interpretation types, is provided in Table 3.2 (as a reference for the subsequent experiments, this is also available in Appendix 2 with some sample answers).

The author and a graduate student majoring in English language education independently categorized 30% of the data, with 90.4% inter-rater agreement. The discrepancies were resolved through discussion, and the author categorized the remaining data.

Table 3.2

Definitions of Interpretation Category its Relationship With Appropriateness According to Semantic Transparency

Category	Definition	Transparent	Opaque
MBI	Literal translations of both two morphemes	○	×
PMI	Translations where part of the morphemes' semantic information was deleted or modified, some information was added, or the interpretation involved their associated meanings. The meanings are semantically appropriate in the contexts.	○	○
CBI	Translations whose meanings were appropriate in the contexts, and no morphemic information was included	○	○
INI	Translations whose meaning was inappropriate in the contexts. It covers cases where part of the morpheme information was included.	×	×
None	No translation corresponding to target words was included.	×	×

As for RQs 1-3, we analyzed how readers could maintain coherent representations of sentences with unknown opaque words by examining the types of information the participants revised. This is because in the opaque condition, the literal understanding of morphemes and context did not make sense in the sentence and some modifications of the textual information, including the target word itself, were necessary to achieve a meaningful interpretation.

First, sentence translation protocols, in which the participants wrote semantically consistent meanings, were extracted. They were then classified according to the types of information the participants revised. Based on the observation of protocols, revision was made

of the following three areas: part of target word, whole target word, and context. Finally, a total of five types were found, including combinations of the two types of revised information.

Table 3.3

Examples of Revision Types in the Translation Protocol in Experiment 1

Revised information	Example
Part of target word	Original: After arriving at the office, he realized that he had left his **pricebattle** on the train. Translation: 会社に到着後、彼は電車の中に財布を忘れたことに気づいた。(After arriving at the office, he realized that he had left his wallet on the train.)
Whole target word	Original: To cook steak for her family, Jessica bought a **lightpole** at the supermarket. Translation: 家族にステーキを焼くために、ジェシカはスーパーで具材を買ってきた。(To cook steak for her family, Jessica bought some ingredients at the supermarket.)
Context	Original: Ellen thought that the woman was a **handknife** from her clothes. Translation: Ellen はその女性が彼女の洋服の中に小さなナイフをしのばせていると思った。(Ellen thought that the woman was concealing a small knife in her clothes.)
Part of target word × Context	Original: Every time Stephanie ate **walkroad** she was happy. Translation: ステファニーは歩いているときいつも幸せそうだ。(Every time Stephanie walked she looked happy.)
Whole target word × Context	Original: When Betty became a university student, she left her **lockerbox** for the first time. Translation: ベティーが大学生になったとき、彼女は初めて厳しい現実を知った。(When Betty became a university student, she realized the harsh reality for the first time.)

Note. Target words are boldfaced. The underlined and double-underlined represent revised target word and context meanings, respectively.

The first two types were those where participants changed the interpretations of target words without changing context interpretations. *Revision of part of target word* means participants revised part of the morphological information, as in PMI. *Revision of whole target*

word refers to cases when they did not include any morphological information of the target words, as in CBI.

Revision of context is where they revised some part of the context information. In addition, they also revised both the target word and context. As a result, *Revision of part of target word* × *context* and *Revision of whole target word* × *context* were added. Examples of the five types of revisions with original sentences (English) and their translations (Japanese) are presented in Table 3.3.

Two raters independently made the initial identifications of meaningful translations and subsequent categorizations of 30% of the data, with interrater agreement rates of 95.2% and 93.1%, respectively. Discrepancies were resolved through discussion, and the author categorized the remaining data.

3.1.3 Results

3.1.3.1 Comprehension questions

The correct answer rates for the comprehension questions in the reading task were sufficiently high (M = 91.92%, SD = 4.60, Min = 81.82%, Max = 100.00%). Thus, it was assumed that all participants had read the contexts for comprehension purposes during the reading task; thus the data from all participants were submitted to the subsequent analysis.

3.1.3.2 Reading times

Table 3.4 provides the descriptive statistics of the reading times for target words in the reading task. It appears that reading times for opaque words and targets in neutral contexts were longer than transparent words and words in more informative contexts, as can be seen in Figure 3.2.

Table 3.4

Descriptive Statistics of Target Word Reading Times (milliseconds) in the Reading Task in Experiment 1

	Informative context			Neutral context		
	M	95% CI	*SD*	*M*	95% CI	*SD*
Transparent	1211.41	[1040.29, 1382.52]	458.26	1419.44	[1206.20, 1632.69]	571.09
Opaque	1419.45	[1190.78, 1648.13]	612.39	1598.50	[1345.38, 1851.62]	677.87

Note. CI = confidence interval.

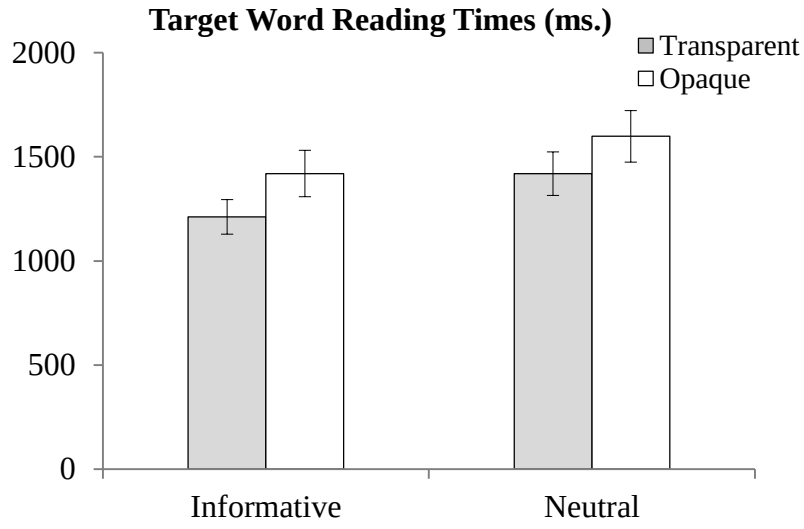


Figure 3.2. Target word reading times in the reading task in Experiment 1.

A 2 (transparency) × 2 (context) ANOVA was conducted (see Table 3.5). The results demonstrated the main effects of transparency and context, where $F(1, 29) = 32.58$, $p < .001$, $\eta_p^2 = .529$, $F(1, 29) = 14.88$, $p = .001$, $\eta_p^2 = .339$, respectively. However, the Transparency × Context interaction was not significant, $F(1, 29) = 0.14$ $p = .713$, $\eta_p^2 = .005$. These results demonstrate that opaque targets took significantly more time to be read than transparent words. This suggests that the participants were sensitive to the morphological and contextual information while reading regardless of the type of contextual information.

Tabel 3.5

Summary Table for Two-Way ANOVA of the Effects of Transparency and Context on Reading Times of Target Words in Experiment 1

Source	*SS*	*df*	*MS*	*F*	*p*	η_p^2
Transparency (T)	1123720.34	1	1123720.34	32.58	< .001	.529
Error (T)	1000204.09	29	34489.80			
Context (C)	1123867.43	1	1123867.43	14.88	.001	.339
Error (C)	2191113.11	29	75555.63			
T × C	6305.76	1	6305.76	0.14	.713	.005
Error (T × C)	1320550.41	29	45536.22			

3.1.3.3 Interpretation of target words

The proportions of appropriate interpretations are presented in Table 3.6. As can be seen, participants made appropriate interpretations for transparent words in almost all cases (Informative: 98.3%; Neutral: 98.3%), while doing so for opaque words in only about half of the cases. In addition, there seems to be no difference by context informativeness.

Table 3.6

Proportions of Appropriate Interpretations in the Translation Task (%) in Experiment 1

	Informative context			Neutral context		
Transparency	*M*	95% CI	*SD*	*M*	95% CI	*SD*
Transparent	98.28	[96.28, 100.27]	5.34	98.28	[95.95, 100.60]	6.24
Opaque	46.98	[38.01, 55.96]	24.04	51.72	[43.20, 60.25]	22.82

To examine the content of their interpretations more closely, Table 3.7 shows the proportion of interpretation types in each condition. Transparent and opaque target words

appear to have their own similar distributions in interpretation irrespective of context informativeness. The participants predominantly made morpheme-based interpretations (MBI).

Table 3.7

Proportions of Each Interpretation Type in the Translations Task (%) in Experiment 1

	Informative context					Neutral context				
Transparency	MBI	PMI	CBI	INI	None	MBI	PMI	CBI	INI	None
Transparent	95.00	2.92	0.42	0.00	2.08	91.67	5.42	1.25	0.42	1.67
	(7.64)	(5.29)	(2.24)	(0.00)	(6.52)	(9.86)	(8.95)	(3.75)	(2.24)	(4.25)
Opaque	48.33	22.08	23.33	2.50	1.25	46.67	35.00	15.00	0.42	2.08
	(28.27)	(16.38)	(17.30)	(6.77)	(3.75)	(23.92)	(18.09)	(13.84)	(2.24)	(4.66)

Note. Standard deviations are in parentheses.

As Figure 3.3 shows, there was greater variance in the interpretation of opaque words than transparent words. However, MBI was the most preferred interpretation type irrespective of context informativeness (Informative: 48.3%; Neutral: 46.7%).

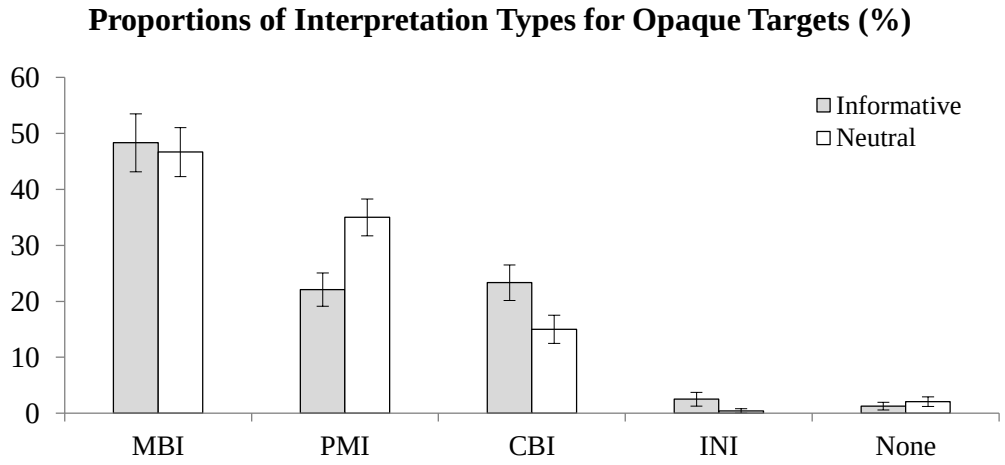


Figure 3.3. Proportion of each interpretation type for opaque target words in Experiment 1.

To explore if there were any differences in the types of information used for interpretation based on context informativeness, a *t*-test compared the proportion of each interpretation category for opaque targets. This analysis was only targeted to opaque words. The results showed that PBI was produced more in neutral contexts than in informative contexts, $t(29) = -4.19$, $p < .001$, $d = 0.73$, while CBI was more common in informative contexts, $t(29) = -2.61$, $p = .014$, $d = 052$. The proportion of MBIs did not differ by Context, $t(29) = 0.46$, $p = .645$, $d = 0.02$.

3.1.3.4 Representation of sentences with opaque unknown words

So far, the participants' interpretations of target words in the translation protocol suggested that readers interpreted opaque words literally (MBI) in about half of the cases. However, it was still possible that the participants had built a coherent representation of the sentence in their mind by changing or revising the interpretations of some information in the surrounding context.

Table 3.8

Frequency of Each Revision Type for Opaque Targets in the Translation Task in Experiment 1

Revised information	Frequency	Percent
Part of target word	123	45.56%
Whole target word	83	30.74%
Context	44	16.30%
Part of target word × Context	17	6.30%
Whole target word × Context	3	1.10%
Total	270	100.00%

Analysis of the sentence translation protocols in the translation task showed that a total of 270 out of 480 protocols (56.3%) were meaningful ones. Table 3.8 provides the frequency of each of the five revision types. As can be seen, some participants revised the context interpretation to achieve meaningful sentence representations (revision of context, part of target word × context, and whole target word × context).

To explore whether there was any difference in proportion of type of revision, a chi-square test was conducted on the frequency of each pattern of revision. The results showed that the frequencies of the types of revision type differed significantly, $\chi^2(4) = 223.89$, $p < .001$, Cramer's $V = .407$. A Bonferroni adjusted multiple comparison revealed that the frequency differed among all the revision types (all *p*s < .01). As a result, the frequencies of types of revision were in the following order: Part of target word (45.6%) > Whole target word (30.7%) > Context (16.3%) > Part of target word × Context (6.3%) > Whole target word × Context (1.1%). This showed a tendency for the participants to revise (a) part of a morpheme more than the whole target word, (b) the target word more than the context, and (c) either the target word or context more than both of them.

3.1.4 Discussion

Are EFL readers sensitive to the semantic relationship between morphological and contextual information during reading? (RQ1-1)

The reading times of target words in the reading task demonstrated that the participants took longer to read opaque target words than transparent ones irrespective of context informativeness. This means that they experienced processing difficulties for opaque words as a result of integrating the semantic information of the morphemes and context (Brusnighan & Folk, 2012). This suggests that EFL readers at the intermediate level are sensitive to the semantic relationship between morphemes and context during reading.

The results are consistent with our prediction, given that past studies showed that such readers could infer the meanings of semitransparent unknown words (Hamada, 2014; Mori & Nagy, 1999). In these studies, successful inference involves perceiving whether morpheme-based meanings are consistent with the context before generating meanings appropriate to the context. What is novel in this study was that this trend was seen in normal reading, where readers' attention is less likely to be directed to each unknown word. It would seem natural given the findings that L2 learners can inconsistent information during reading that contradicted with prior descriptions (Morishima, 2013; Ushiro, Nahatame, et al., 2016).

However, readers' awareness in neutral contexts was not fond in L1 reading (Brusnighan & Folk, 2012, Experiment 1). The discrepancy appears strange because the proficiencies of L1 readers are much higher than those of the present participants. In trying to understand this problem, one might argue that the neutral contexts provided ample information about the target words and thus behaved like informative contexts. However, the main effect of Context suggests that both transparent and opaque targets in neutral contexts received longer reading times than those in informative contexts; this observation was also found in the case of unknown monomorphemic words (Hamada, 2013). Thus, it is unlikely that neutral contexts behaved informatively.

A possible explanation is that the semantic inconsistency of opaque words in the neutral context was more salient in this study. In Brusnighan and Folk (2012), novel opaque words were created by compounding morphemes whose literal meanings were inconsistent with the original words, not necessarily with the meanings of the surrounding contexts. In their study, for example, the novel opaque word *deskdoor* was not helpful in determining the meaning of the comparable existing transparent word *milkshake*, and similarly with its novel counterpart *drinkblend*. As a result, in the informative context *The party host used a blender to mix each guest a deskdoor last night*, the semantic inconsistency was salient because the sentence context

was strongly biased to the meaning of transparent words (i.e., *cocktail*, *drinkblend*). On the contrary, in the neutral context *Matthew got off work and bought a deskdoor afterword*, the literal meaning of the target word, as well as transparent one, is possible, in that both meanings are related to something we can buy. Thus, this material characteristics of the neutral context might have resulted in the lack of a difference between the transparent and opaque conditions.

In the present experiment, however, the semantic transparency of the target words was manipulated by the semantic in/consistency between literal word-based interpretation and context meanings regardless of context informativeness, as confirmed in the norming study (semantic fitness judgment). This was because we were interested in whether EFL readers could consider both morphemes and context in processing and interpretation, and it has often been reported that readers make literal interpretations even when context meanings do not support them (e.g., Laufer, 1989). Thus, in the present study, morpheme-based interpretations of opaque words were always semantically inconsistent even in a neutral context.

For all the differences in material characteristic between the two studies, it is important that EFL readers could integrate the morphological information with context meaning even when it was abstract, suggesting their strong sensitivity to the semantic relation between the two sources during reading.

What types of information do intermediate-level Japanese EFL readers use to interpret transparent and opaque unknown words? (RQ1-2)

The analysis of target word interpretations in the translation task showed that participants relied on different sources of information depending on the transparency of morphemes. For transparent words, they interpreted them predominantly based on the morphemes. The convergence of the two sources of information led them to use morphological information as having more concrete semantics than the contextual meaning. This result is in line with previous

studies employing the multiple-choice format (Brusnighan & Folk, 2012, Experiment 2; Hamada, 2014).

However, this reliance on morphemes was also seen for opaque words, for which word-based, literal interpretation contradicted the surrounding context meanings. Such cases accounted for approximately half of the cases (Informative: 48.3%; Neutral: 46.7%). Given the results of the reading time data, the interpretation trend is inconsistent with previous studies that concluded that misinterpretation based on word information comes from lack of monitoring (Hamada, 2014). According to this study, the primary source of incorrect word-based inferences was their failure to notice the semantic inconsistency.

This discrepancy can be related to the methodological difference between the two experiments. Previous studies used a multiple-choice format, which gave options presenting different sources of information used for interpretation (e.g., morphemes, context). Therefore, once the participants perceived the semantic inconsistency, the presence of answer options would allow them to reject the literal option and explore other possible meanings more easily. In contrast, the absence of such options in this experiment might have made it difficult to go beyond literal interpretations because the participants had to generate contextually appropriate meanings on their own, at the same time suppressing word-based information. This kind of difficulty in flexibly changing the interpretation was also seen in the investigation with intermediate-level learners (Ushiro et al., 2010), where their performance was much superior in the multiple-choice task (90.7%) than in the open-ended task (50.2%).

Combined with the reading time data, it is suggested that the perception of inconsistency does not always result in deriving a contextually appropriate meaning, and this discrepancy would come from the gap between the two processes: the meaning generation process is more demanding than that of perceiving the relationship. This affords us new insight into the causes of the wrong word-based interpretation.

As for the appropriate interpretations for opaque target words, two types of interpretations were assigned in this study, PMIs and CBIs. The two interpretations were both contextually appropriate but differed in whether or not they included morphological information in their interpretations. Here, it was found that the relative reliance on morphemes and contexts can be affected by the context quality. The *t*-tests showed that participants made more PMIs for neutral contexts (Informative: 22.1%; Neutral: 35.0%) while CBIs were more commonly made in informative contexts (Informative: 23.3%; Neutral: 15.0%).

The results can be explained in terms of the amount of information constructed from the context for target words. In informative contexts, readers can activate or narrow down more concretely the semantics of upcoming words; thus, they were more likely to suppress morphemic information contradicting their prediction and prioritize contextual information. In contrast, readers could derive only abstract semantic information from neutral contexts. Thus, they had to include or rely on morphological information, since these words (noun-based morphemes) had some concrete meanings. This is in line with the findings that biasing context will affect lexical access (Duffy et al., 1988; Elston-Güttler & Friederici, 2005), suppressing unnecessary meanings (Ushiro et al., 2013). Although there was no superiority or inferiority of the two interpretation types in this study, it can be argued that the amount of information constructed from context can affect how much readers base their interpretations on morphological or contextual information.

What types of information do intermediate-level Japanese EFL readers revise to achieve coherent representations of a sentence with opaque unknown words? (RQ1-3)

In order to examine how they achieve coherent representation of the context sentences, the translation protocols for the opaque condition were analyzed. The results showed that in 270 out of 480 cases (56.3%), the participants made semantically coherent representations by

revising literal interpretations of some part of the sentence. The revision was made not only to a part or the whole of the target word but also to the context, as well as to both target word and context.

The comparison demonstrated that the frequency with which readers made revisions differed by the types of revised information as follows: Part of target word (45.6%) > Whole target word (30.7%) > Context (16.3%) > Part of target word × Context (6.3%) > Whole target word × Context (1.1%). This order of revision should reflect some tendencies or nature of readers' revision, which will be discussed below.

First, readers would prefer to revise the interpretation of the constituent morphemes of the unknown words. This is quite natural given that although the target words consisted of familiar words, the words as a whole were unfamiliar to them. Thus, they found their literal interpretations unreliable and rejected a part or all of the morphological information to explore context-based meanings. In addition, the higher proportion of revision of Part of the target word over the Whole target word might reflect a preference for using morphological information in their interpretation as much as possible. This is also consistent with previous findings that L2 learners, with some proficiency, tried to integrate both morphemes and contexts rather than using context only (Mori, 2002; Mori & Nagy, 1999).

Secondly, readers sometimes changed the interpretations of contextual sentences whose constituent words were familiar to them. This interpretation can be regarded as distortion rather than pure revision because the running words in the contextual sentence were all high-frequency known words. Still, there is a possibility that readers had assumed the context word was used in a novel meaning and retained the literal meaning of the target word. However, given the fact the target words, though consisting of known words (morphemes), were unfamiliar as a whole, changing the interpretation of the context would be undesirable.

Similar distortions of the surrounding context were reported in an earlier study (Bensoussan & Laufer, 1984). However, an important difference is that in their study, distortion took place *after* inferring an unknown word. The participant interpreted the subsequent context so that the wrongly inferred meanings would make sense. In addition, such distortions were the result of readers' misidentification of unknown words as known words. On the other hand, the cases found in this study were possibly made in the interpretation process after participants noticed the semantic inconsistency between the two sources. This kind of modification would reflect their mental effort to achieve meaningful understanding, but this would also lead to the inaccurate understanding of a text. However, this possibility will be closely examined by investigating learners' inferential processes in the think-aloud study (Experiment 3).

The results of this study would indicate the importance of examining and confirming the interpretation of not only unknown words but also the surrounding context. This is because in such cases, even when readers could make appropriate interpretations of unknown words, their text understanding was affected by changing the interpretation of context meanings.

3.1.5 Conclusion of Experiment 1

The purposes of Experiment 1 were to investigate whether or not EFL readers were able to interpret unknown words in reading, and identify the source of their difficulties, if any, in this process—that is, the perception of semantic in/consistency between the two sources—and the generation of contextually appropriate meanings. Moreover, the way in which they tried to achieve a coherent representation of the sentence was explored. To this end, the reading task examined the on-line processing of the semantic integration of the two sources, and translation protocols of both the target word alone and the entire sentence were analyzed.

The reading time results showed that they exhibited a stable sensitivity to the relation between morphological and contextual information during reading even in a neutral context,

where only the abstract semantics was available. As for interpretation, however, they often made incorrect morpheme-based interpretations of opaque words that were inconsistent with the contextual meanings. Thus, EFL readers, even at the intermediate level, had difficulty interpreting unknown words based on the semantic transparency of the words. Combined with the online process, the particular source of difficulty lies in the meaning generation process after perceiving the semantic relationship. In addition, incorrect word-based interpretations sometimes would make them distort the surrounding context. In that sense, failure in word-level inference could affect the comprehension of the wider context, as well as the unknown target words.

3.2 Experiment 2: Semantic Memory Representation of Unknown Words and its Relationship With Text Processing

3.2.1 Purposes, overview, and research questions of Experiment 2

Experiment 1 showed the challenges of interpreting opaque words and their negative impacts on context interpretations. However, there are some limitations that should be addressed in Experiment 2 to definitively conclude how such words affect reading comprehension, as well as word interpretation.

First, Experiment 1 employed the translation task because that allowed us to examine the interpretation of both the target words and their surrounding context at the same time. However, there was a possibility that the task itself might have solicited a different processing of the target words from that in normal reading. The task required participants to produce a complete sentence translation; therefore, the presence of unknown words made them infer their meanings for task completion. In contrast, the importance of each unknown word during normal reading is relatively low, as can be seen in the high proportion of ignored unknown words (Fraser, 1999; Paribakht & Wesche, 1999). Although translating a whole sentence made unknown words less

salient than in tasks explicitly requiring words meanings (e.g., lexical inferencing task), it is highly possible that they were more motivated to reveal the meanings of target words more than in normal reading. For this reason, Experiment 2 examined the interpretation of target words in a more indirect way. Specifically, we investigated what semantic memory representations readers had for target words as a result of reading. This was realized by asking participants to answer comprehension questions whose answers were related to the meaning of the target word.

Another limitation was related to the text length. Experiment 1 used a single sentence context with either informative or neutral information about the target word. Therefore, one could argue that it is unnatural for readers to read such a short text, and the results therefore underestimated their performance because discourse-based information was unavailable. Indeed, some participants in Experiment 1 reported the lack of context to determine specific meanings as a problem, and as a result, felt uncertain about which sources to rely on when interpreting the target words. Therefore, it was possible that readers could avoid such overreliance on morphological information if additional context was available.

Therefore, the present experiment used passages consisting of two sentences: first sentences with either transparent or opaque target words as in Experiment 1, and second sentences with informative information for target word meanings. Though just two sentences still do not constitute naturalistic texts, the results of the present experiment can to some extent be generalized to how readers process and interpret unknown words in reading for the following reasons. First, although discourse-level information could narrow down more exact meanings of the word in a given text if used effectively, such clues beyond the sentence boundary are less likely to be used due to limited cognitive capacity especially in reading (Pulido, 2009; Wesche & Paribakht, 2010). Second, the use of directive context (second sentence) would qualitatively complement the lack of quantity (length) of contextual information. Since most naturalistic contexts do not provide sufficient information to determine a word's meaning (Hulstijn, 1992;

Beck et al., 1983), the addition of one subsequent rich context, irrespective of the context quality of the first sentence, provides sufficient semantic clues, similar to what would be derived from a longer text, though the context quality within a single sentence is still considered to be a critical factor in determining the success of inference (Wesche & Paribakht, 2010).

In the experiment, participants read the two-sentence passages, the first sentence containing target words (transparent or opaque) with either informative or neutral contextual clues, and the second sentence with informative clues for inferring the context-based meanings of the target words. After reading the passages, they answered comprehension questions asking the meanings of target words. As in Experiment 1, the reading times of the first and second sentences in transparent and opaque conditions were compared to examine their on-line processing of such factors as the perception of the semantic relation between the two sources and the mental effort of inferring the word meaning (Oakhill et al., 2016). The answers to the comprehension questions were analyzed to see what types of semantic information (morpheme, context) were used for interpretation. These points are summarized in the following research questions.

RQ 2-1: How do intermediate-level Japanese EFL readers process passages with transparent and opaque unknown words?

RQ 2-2: What kinds of semantic representations do intermediate-level Japanese EFL readers construct for unknown words through reading?

The findings of this experiment, combined with those of Experiment 1, would have some implications for how to introduce new words and have students deal with them in reading instruction. If readers could interpret unknown words appropriately, the effects of inappropriate word-based interpretation on reading comprehension would be relatively local because readers

would suffer misinterpretation only when encountering unknown words (in the first sentence), but later could modify it appropriately in the subsequent context (in the second sentence). On the other hand, if they could not modify their initial understanding and interpret the text so that the word-based interpretation makes sense (Bensoussan & Laufer, 1984), then text comprehension would be affected more globally. This information tells us how and when teachers should support or intervene in students' understanding.

3.2.2 Method

3.2.2.1 Participants

The participants in Experiment 2 were 24 university undergraduate and graduate students from various majors. Every student had learned EFL in Japan for more than six years. Their estimated English proficiency ranged from intermediate to more advanced levels based on self-reported scores on English proficiency tests, such as the TOEIC Listening and Reading test (n = 13, M = 751.92, SD = 126.59, range = 495–905), TOEFL ITP test (n = 1, M = 580), and EIKEN test (3rd grade: n = 1; pre-2nd grade: n = 5; 2nd grade, n = 6; pre-1st grade: n = 4). Their reading proficiency was tested with the reading proficiency test used in Experiment 1. The average score was 16.00 (SD = 3.38), ranging from 12 to 24 (Cronbach's α = .75).

3.2.2.2 Materials

Experimental passages and comprehension questions

In this experiment, the participants were asked to read two-sentence passages in which either transparent or opaque targets were embedded in the first sentence with different context informativeness (i.e., informative vs. neutral). The second sentence always supported the contextually appropriate meanings of target words presented in the first sentence. Although the

target words were the same as in the other experiments, the contexts were newly created for this experiment. Below is the detailed procedure of context creation.

First, each of the 32 target words was embedded in two sentence-context passages consisting of a first and second sentence. The first sentences conveyed either informative or neutral context information about target word meanings. In the informative context, the concrete meanings of the target words could be inferred from the surrounding context, while abstract meanings could be inferred from the neutral context, as in Experiment 1. In order to make the comprehension questions identical regardless of context type, both informative and neutral contexts shared common constituents, such as the name of the protagonist, and phrases including the target words. For example, in the context sentences for the target word *businessbag*, the informative first sentence was *William realized he had left his businessbag when he arrived at the office.* while the neutral counterpart was *William realized that he had left his businessbag somewhere*. In both contexts, readers could infer that the protagonist (i.e., William) had left something, but in the informative context, readers could narrow down the meaning to something that one normally brings to the office, while it could only be narrowed down to some belonging in the neutral context. The target words in the above contexts were regarded as transparent words, as the morpheme-based meanings were consistent with their surrounding context. In order to create the opaque targets, the 32 target words were regrouped into 16 pairs of target words. Then, the contextual sentence for one target word was exchanged for the other target word so that the contexts used for one target were inconsistent with the contexts used for the other target.

The second sentence was identical irrespective of context informativeness and the semantic transparency of target words. All second sentences presented information conducive to inferring the target word meaning. For the above example of *businessbag*, the second sentence was *He was upset because his computer was in it*. Combined with the meaning of the

first sentence, the second sentence conveyed information that contributed to inferring the actual meaning of the target word, in this case something used to put one's belongings in. The criterion for context quality was that used for the informative context in the first sentence. Therefore, the semantic information that could be derived from the two sentences was almost the same regardless of the context quality of the first sentence. In addition, readers would not be able to narrow it down to a single word so as to distinguish which information (i.e., morphemes, contexts) readers used to interpret the target words.

As in Experiment 1, the topic of the passages was daily life, and the context mostly consisted of high-frequency words (JACET 8000, Level 3 or below). The context features of the experimental passages (i.e., number of words and syllables) are presented in Table 3.9. The number of syllables of pseudo target words were calculated based on the sum of the number of syllables of the two morphemes.

Table 3.9

Number of Words and Syllables in Target Sentences in Experiment 2

	Informative context				Neutral context			
	Words		Syllables		Words		Syllables	
Transparency	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
				First sentence				
Transparent	11.80	1.95	15.53	2.74	9.03	1.62	11.71	2.29
Opaque	11.68	1.97	15.72	2.58	9.09	1.62	11.92	2.23
				Second sentence				
Transparent	9.78	0.90	12.22	1.29	9.78	0.70	12.22	1.33
Opaque	9.78	0.70	12.19	1.37	9.78	0.70	12.19	1.37

To tap into what semantic representation readers had for the target word as a result of reading, comprehension questions whose possible answers were phrases containing the target word were created. Therefore, the answers provided by participants would reflect the semantic representations of target words in the process of reading comprehension.

All questions were in the form of *wh*-interrogatives and their possible answers corresponded to noun phrases or prepositional phrases (preposition + noun phrases) that included target words. The question statements contained as minimal amount of information as possible so that participants would not reconsider context sentences based on the information in the question statements. Most of the questions consisted of "interrogative + did + subject + verb?" structure (e.g., *What did William leave?*). Table 3.10 presents an example of experimental passages and their corresponding comprehension questions.

An answer sheet was also prepared for them to write down their answers to the comprehension questions. On this sheet, the question number and space to write down their answers were provided. Participants were required to answer questions in their L1 (Japanese) because if it were made in the target language (English), it would not be clear whether the meaning they obtained was the literal one (or part of it) or based on the context. Therefore, answering in Japanese would be more suitable to tap into the semantic representations of target words.

It should be noted, however, that these questions were not able to directly examine the semantic representations readers had for target words *during* reading because the participants answered the questions *after* reading each passage and were able to think retrospectively. However, since the primary purpose of this experiment was to examine the semantic representations of target words as a part of reading comprehension or text memory, the timing of answering questions was not considered problematic.

Table 3.10

Example of a Pair of Experimental Passages and Comprehension Questions in Experiment 2

S1	Informative:	The shop reduced its production costs because it experienced a **pricebattle** / **businessbag** in the area.	William realized that he had left his **businessbag** / **pricebattle** when he arrived at the office.
	Neutral:	The shop experienced a **pricebattle** / **businessbag** in the area.	William realized that he had left his **businessbag** / **pricebattle** somewhere.
S2:		It was so severe that the shop quickly shut down.	He was upset because his computer was in it.
CQ:		What did the shop experience?	What did William leave?

Note. The first and second target words (boldface) served as transparent and opaque target words, respectively. S1 = first sentence, and S2 = second sentence.

Filler passages

In order to distract readers from the purpose of the experiment, 32 filler passages, also consisting of two sentences, were created based on past studies (e.g., Brusnighan & Folk, 2012). All the filler passages were paired with comprehension questions. Of these, 15 passages contained their possible answers in the first sentence and 17 passages in the second sentence. This was to motivate participants to carefully read both sentences to answer the questions because possible answer phrases for experimental passages were always located in the first sentences. As a practice session, an additional three passages were created with identical passage structures. A complete list of practice, experimental, filler passages are available in Appendix 4.

Norming study

In order to check the (a) validity of the amount of contextual information (informativeness judgment), (b) semantic transparency of the target words (semantic fitness

judgment), and (c) appropriateness of the comprehension question (appropriateness verification), a norming study was conducted. The participants of the norming study were two graduate students majoring in English language education.

The semantic fitness judgment was made to validate the semantic transparency/opacity of target words in relation to their surrounding contextual meanings. The raters read the first sentence and judged whether or not the morpheme-based meanings of target words semantically fit the context in order to confirm that transparent words would receive "Yes" (semantically plausible) versus "No" (semantically implausible) for opaque words.

Next, they made an informativeness judgment**.** In this task, the experimental passages were presented with the target words omitted, and the raters judged whether the context included the intended information in each of the first sentences and in both the first and second sentences. For the first sentences, the raters judged whether concrete semantics could be derived from informative contexts, even though enabling several possible meanings, while only abstract meanings or general categories could be derived for the neutral context. In addition, they were asked to judge whether the cumulative meanings of the two sentences provided informative information about the target words. This was to ensure that the participants could infer the concrete semantics of the target words after reading the passages even when the first sentence was neutral context.

Finally, the validity of the comprehension question was examined. The raters checked whether the possible answers for the comprehension questions corresponded to the target words (replaced with blanks). For filler passages, potential answers were underlined. Although some articles and pronouns (e.g., *a, his*) ideally should be included in the answer, this was neglected because participants would answer in Japanese in the main experiment, and these qualifiers are not necessarily reflected in Japanese when they are inferable from the surrounding contexts.

They were asked to check if the blank would be the only answer for the question by choosing "Yes" or "No."

After finishing all of the judgments, most of the items received the intended responses from the two raters: 92.1% (transparency rating), 94.8% (informativeness rating [first sentence]), 97.6% (informativeness rating [second sentence]), and 91.5% (appropriateness judgment). The items that received unintended responses from at least one rater were revised following their suggestions. After that, the English expression was checked with the help of a native speaker of English, then four presentation sets were created for the experiment. In each set, each of the 32 target words appeared in any of the four conditions, crossing 2 (context) × 2 (transparency).

3.2.2.3 Procedure

The experiment was conducted individually in a silent room. First the author explained the general purpose and entire procedure of the experiment. They were then assigned to one of the four presentation sets for the reading task.

The reading task was administered using SuperLab 5.0 and Response Pad RB-740. The first sentence, second sentence, and comprehension question were presented one at a time (non-cumulative presentation) and their reading times (ms.) were recorded. Figure 3.4 graphically shows the sequence of a trial in the reading task.

First, the participants were given written instructions on the overall procedure of how the task would proceed. They were instructed to read two-sentence passages sentence by sentence on a PC screen, then write down the answer to the comprehension question in Japanese in the answer sheet. They were instructed to read the passages for comprehension.

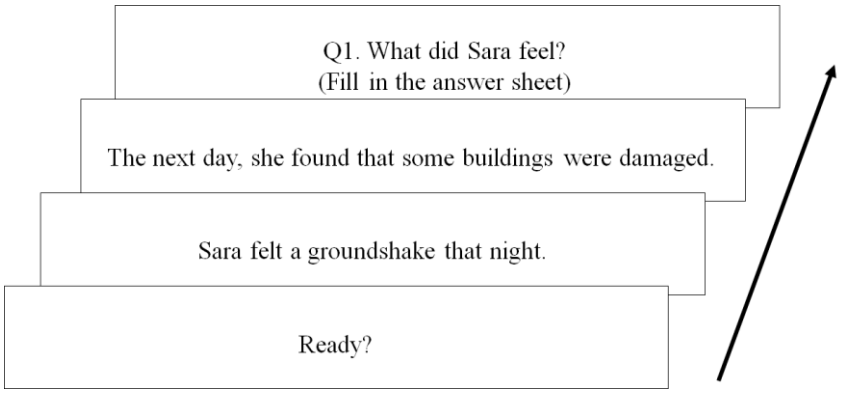


Figure 3.4. An example of the reading task in Experiment 2.

Each trial began with a "Ready?" screen. When participants were ready, they were asked to push the "Yes" button. The next screen presented the first sentence and they were asked to push the button when they thought they had comprehended the sentence. The next button press made the first sentence disappear and the second sentence appear. Similarly, when they finished reading the sentence and pressed the button, the comprehension question appeared in the center of the screen. At that time, the participants were asked to write down their answers in Japanese on the answer sheet. The question numbers were always presented on the comprehension question screen so that they could keep track of it. They were encouraged to answer all questions, but if they could not come up with an answer, they were allowed to write 分からない (*I do not know*). Also, when they skipped any of context sentences or a comprehension question, they were instructed to write スキップ (*skip*).

To familiarize participants with this procedure, they performed three practice trials before the main session. The passages and comprehension questions used for the practice had the same structure as the filler passages. The practice trial was followed by the main trials, in which the participants read the 64 passages and answered the accompanying questions after each passage. They were allowed to take a break whenever they needed it. The reading task lasted approximately 25 to 45 minutes. When the participants finished the task, they were informed

that target words were not real English words. After the reading task and a short break, they took the 20-minute reading proficiency test.

3.2.2.4 Scoring and analysis

Reading times

Participants' reading times on the first and second sentences of the reading task were analyzed to examine their sensitivity to the two sources and also their mental effort in inferring the context-based meaning. This was done by comparing the readings times on the transparency and opaque conditions.

First, in order to offset the effects of sentence lengths, the reading times (ms.) were divided by the number of syllables of each sentence. In so doing, the syllables of the target pseudo words were calculated by the sum of each of the two morphemes (nouns) because, given the ease of decomposition, they were expected to be read like two words.

Next, reading time outliers were treated. First of all, reading times above 3*SD*s of each participant reading times per condition were regarded as outliers and discarded. Also, the trials in which participants reported that they skipped any of the first or second sentences or comprehension questions, or reading times below 200ms were removed as operational mistakes. As a result, 4.8% of the trials were discarded, and the remaining data were submitted to the following analysis.

To examine RQ1, an analysis of reading times was conducted for both first and second sentences. A 2 (transparency: transparent, opaque) × 2 (context: informative, neutral) × 2 (sentence: first sentence, second sentence) three-way ANOVA was run. The three independent variables were all within-participant designs.

In interpreting the reading time data, we considered that what longer reading times for the opaque condition than the transparent condition indicate could differ according to the

section of the passage (i.e., first or second sentence). For the first sentence, much of the longer reading times for opaque conditions relative to transparent conditions reflects readers' perception and reaction to the inconsistency conveyed by the two sources, though they sometimes might reflect the mental effort to make the inference (Oakhill et al., 2016). For the second sentence, the difference would mostly reflect the mental effort in inferring appropriate meanings, because readers would not spend more time reading unless they had detected the inconsistencies while reading the first sentences. This issue will be discussed in detail in the discussion section.

Comprehension questions

In order to examine what semantic representations readers had for the target words, participants' answers to the comprehension questions were classified into five types (MBI: morpheme-based interpretation; PMI: partially-morpheme-based interpretation; CBI: context-based interpretation; INI: inappropriate interpretation; None [*I don't know*]) as in Experiment 1 (for the definitions, see Appendix 2). Again, for transparent words, MBIs, PBIs, and CBIs were considered appropriate, whereas PBIs and CBIs were appropriate for opaque targets. Two raters independently categorized 30% of the data, with 91.9% agreement. The discrepancies were resolved through discussion, and the author categorized the remaining data.

3.2.3 Results

3.2.3.1 Reading times of first and second sentences

Descriptive statistics of the reading times in the first and second sentences are presented in Table 3.11 and graphically in Figure 3.5. Apparently, reading times were longer for the opaque condition regardless of Context or Sentence.

Table 3.11

Descriptive Statistics of Target Sentence Reading Times (per Syllable) in Each Sentence in Experiment 2

	Informative context			Neutral context		
Transparency	*M*	95% CI	*SD*	*M*	95% CI	*SD*
			First sentence			
Transparent	564.60	[500.74, 628.45]	151.23	654.99	[576.64, 733.34]	185.55
Opaque	828.71	[743.92, 913.50]	200.80	921.19	[801.07, 1041.31]	284.47
			Second sentence			
Transparent	365.88	[321.77, 409.98]	104.46	375.02	[335.93, 414.11]	92.57
Opaque	459.42	[406.10, 512.74]	126.27	550.31	[482.77, 617.84]	159.94

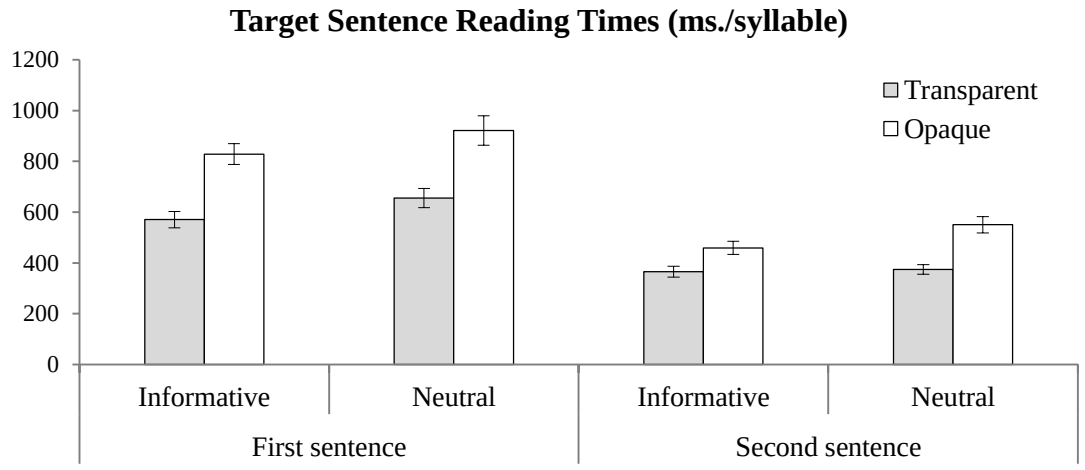


Figure 3.5. Reading times for the first and second sentences in Experiment 2.

A (transparency) × (context) × (sentence) three-way ANOVA was conducted on the reading times of the first and second sentences (see Table 3.12). The results showed that the main effects of transparency, context, and sentence were all statistically significant, $F(1, 23) = 153.13$, $p < .001$, $\eta_p^2 = .869$; $F(1, 23) = 33.57$, $p < .001$, $\eta_p^2 = .593$; $F(1, 23) = 108.76$, $p < .001$; $\eta_p^2 = .825$, respectively.

More importantly, the three-way interaction reached marginal significance, $F(1, 23) = 3.04$, $p = .094$, $\eta_p^2 = .117$. To interpret this interaction, follow-up analysis was conducted. Because of our interests, this was only targeted to reveal the effects of transparency and context on reading times within each of the first and second sentences.

A post-hock test showed that the Transparency × Sentence two-way interaction was significant in both informative ($p < .001$) and neutral contexts ($p = .015$). In addition, the Transparency × Context two-way interaction was also significant for the second sentence ($p = .013$). A follow-up test showed that reading times for transparent conditions differed by context in the first sentence ($p = .001$), but not in the second sentence ($p = .739$).

In sum, these results showed that the reading times for opaque conditions were longer than for transparent words irrespective of context informativeness and sentence position. In addition, reading times were longer for neutral conditions than for informative conditions, except for transparent conditions of the second sentence.

Table 3.12

Summary Table for Three-Way ANOVA of the Effects of Transparency, Context, and Sentence on Reading Times in Experiment 2

Source	*SS*	*df*	*MS*	*F*	*p*	η_p^2
Transparency (T)	1915874.65	1	1915874.65	153.13	< .001	.869
Error (T)	287761.31	23	12511.36			
Context (C)	240101.37	1	240101.37	33.57	< .001	.593
Error (C)	164502.00	23	7152.26			
Sentence (S)	4456833.80	1	4456833.80	108.76	< .001	.825
Error (S)	942476.99	23	40977.26			
T × C	21080.63	1	21080.63	3.56	.072	.134
Error (T × C)	136181.43	23	5920.93			
T × S	205106.93	1	205106.94	21.50	< .001	.483
Error (T × S)	219386.15	23	9538.53			
C × S	20590.84	1	20590.84	1.23	.278	.051
Error (C × S)	384028.45	23	16696.89			
T × C × S	19036.20	1	19036.20	3.04	.094	.117

3.2.3.2 Interpretations of target words

Table 3.12 provides the descriptive statistics of the proportions of appropriate interpretations observed in the answers to the comprehension questions. For transparent words, the participants made appropriate interpretations in almost all cases. On the other hand, the proportions were below 50% in both informative (42.3%) and neutral (33.0%) context conditions for opaque words. However, it should be noted that there was a large individual difference in the proportion: M = 37.69%, SD = 24.15, Min = 0.00%, Max = 87.50%. Moreover, of the 24 participants, the proportions for six participants were above mean+1SD (61.84%) while five participants were below mean−1SD (13.54%).

Table 3.13

Proportion of Appropriate Interpretations (%) in Experiment 2

	Informative context			Neutral context		
Transparency	*M*	95% CI	*SD*	*M*	95% CI	*SD*
Transparent	98.44	[96.75, 100.13]	4.22	96.35	[94.03, 98.68]	5.80
Opaque	42.34	[31.48, 53.19]	27.13	33.04	[22.09, 43.98]	27.35

A 2 (transparency) × 2 (context) two-way ANOVA was performed (see Table 3.14). The results showed that the main effects of transparency, $F(1, 23) = 140.49$, $p < .001$, $\eta_p^2 = .859$, and context, $F(1, 23) = 4.45$, $p = .046$, $\eta_p^2 = .162$, were significant. However, the Transparency × Context interaction did not reach significance ($p = .177$). It should be noted that the apparent ceiling effect observed for transparent words, the source of the main effect of context, would be caused mainly by the difference in opaque words. These results demonstrate that (a) the proportions of appropriate interpretations were higher for transparent conditions, and (b) those for opaque words were higher when the first sentence was in an informative context.

Tabel 3.14

Summary Table for Two-Way ANOVA of the Effects of Transparency and Context on Proportions of Appropriate Interpretations of Target Words in Experiment 2

Source	*SS*	*df*	*MS*	*F*	*p*	η_p^2
Transparency (T)	85566.22	1	85566.22	140.49	< .001	.859
Error (T)	14008.69	23	609.07			
Context (C)	777.54	1	777.54	4.45	.046	.162
Error (C)	4020.75	23	174.82			
T × C	312.52	1	312.52	1.94	.177	.078
Error (T × C)	3704.52	23	161.07			

Table 3.15 provides the proportions of the interpretation type in each condition. As in Experiment 1, MBIs were predominantly made for transparent words (about 90%). With regard to opaque words, there were variations of interpretations (see also Figure 3.6). As for inappropriate interpretations, MBI (Informative: 29.2%; Neutral: 40.9%) and None (Informative: 26.7%; Neutral: 25.0%) were found frequently, with little INI. As can be seen in the standard deviations, there was great variation in the interpretation type.

Table 3.15

Proportions of Each Interpretation Type in the Reading Task (%) in Experiment 2

	Informative context					Neutral context				
Transparency	MBI	PMI	CBI	INI	None	MBI	PMI	CBI	INI	None
Transparent	89.76	8.16	0.52	0.00	1.56	90.10	5.21	1.04	1.04	2.81
	(11.65)	(10.39)	(2.55)	(0.00)	(4.22)	(10.41)	(8.17)	(3.53)	(3.53)	(9.62)
Opaque	29.17	13.69	28.65	1.64	26.94	40.92	13.54	19.49	1.04	25.00
	(31.84)	(13.75)	(21.26)	(4.43)	(30.71)	(32.31)	(18.77)	(17.94)	(3.53)	(32.13)

Note. Standard deviations are in parentheses.

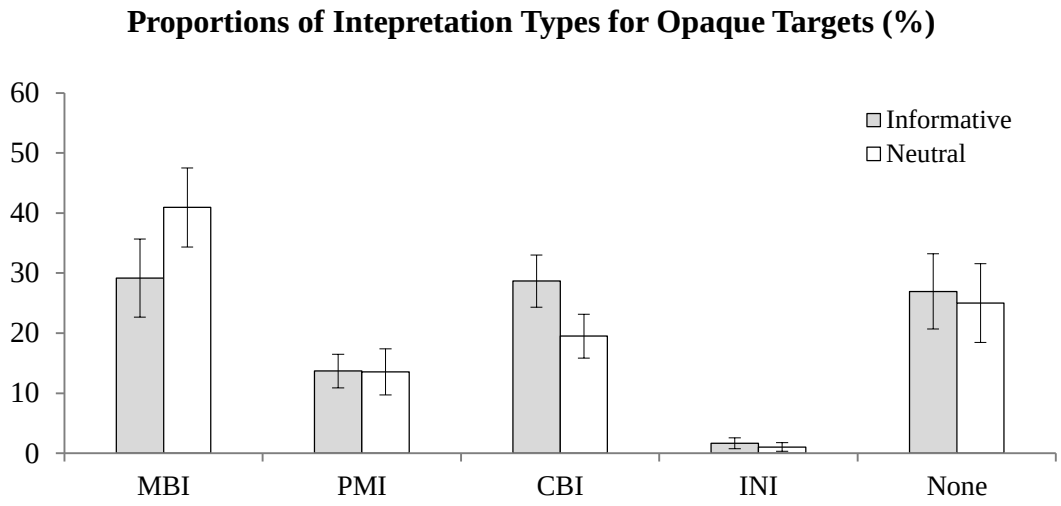


Figure 3.6. Proportion of each interpretation type for opaque words in Experiment 2.

To examine the effects of the context informativeness of first sentences on interpretation types for opaque targets, a paired *t*-test was conducted. The results showed that MBIs were more often made in neutral contexts than in informative ones, $t(23) = -2.82$, $p = .010$, $d = .367$, while CBIs were more often produced in informative contexts, $t(23) = 2.25$, $p = .034$, $d = .455$. For the other interpretation types, significant difference was not observed as a function of context (all *p*s > .10).

3.2.4 Discussion

How do intermediate-level Japanese EFL readers process sentences with transparent and opaque unknown words? (RQ2-1)

On-line reading time data were analyzed to explore how readers processed a text with unknown words, especially opaque words. Overall, the reading times were longer for opaque conditions irrespective of context and sentence. This means that participants were sensitive to the semantic relationship between morphemes and context during reading (Brusnighan & Folk, 2012). Interestingly, the three-way interaction showed the possibility that processing behavior

was different between the first and second sentence. Therefore, the discussion here focuses on what cognitive behavior they engaged while reading in each of the first and second sentences.

Processing of the first sentences. The reading time results for the first sentences showed that participants spent more time reading them when they contained opaque words than transparent ones. This means that they experienced processing difficulties integrating semantic information from the morphemes and context (Brusnighan & Folk, 2012; Oakhill et al., 2016). As with Experiment 1, this trend was also seen in neutral contexts, suggesting that they could use abstract information derived from the neutral sentence to monitor the semantic relations.

However, what cognitive processing the participants were engaging in during the inflated reading times needs careful consideration, i.e., trying to infer the meanings of the target words or only reacting to the processing difficulty (Oakhill et al., 2016). Originally, it was presumed that if they had attempted to infer the context-based meaning, more mental effort would have been required in the neutral contexts due to the smaller amount of information derived from the context, and they indeed spent more time reading the neutral contexts in the opaque condition (M_{diff} = 92.5ms). However, the lack of Transparency × Context interaction suggests that longer reading times in neutral contexts were also seen in the transparent condition (M_{diff} = 90.4ms), where there were no semantic conflicts. In other words, the neutral context itself elicited longer reading times irrespective of the opacity of the target words. For this reason, it is difficult to come to a clear conclusion as to their inferential process after noticing the semantic inconsistency based only on these findings; thus, we will discuss this issue later in conjunction with other data.

Processing of the second sentences. Now let us consider the reading processes in the second sentences. The second sentences always contained informative clues about the target words presented in the first sentences. The reading time results showed that the participants spent more time reading the second sentence when the first sentence had opaque unknown

words regardless of the context quality of the first sentence. Furthermore, unlike the first sentence, the longer reading times for the neutral context were observed only for the opaque condition (M_{diff} = 90.9ms), but not for the transparent condition (M_{diff} = 9.1ms).

For transparent words, the lack of a difference in reading times by context means that the readers did not feel difficulty processing the sentence. This is because if they had not resolved the unknown words when they had finished reading the first sentence, more processing costs in the second sentence was required when the first sentence was neutral context. Thus, readers would have been able to build a consistent representation of the text when they had finished the first sentence due to the convergence of the two sources. Therefore, despite the difference in context quality, their reading times in the second sentence did not differ when the preceding context contained unknown transparent words.

On the other hand, longer reading times for the opaque conditions relative to transparent ones suggest that they kept trying to resolve the semantic conflicts introduced in the first sentence. In addition, the amount of mental effort differed according to the context quality of the first sentence: the reading times for opaque words were longer when the first sentence was a neutral context than an informative one. After reading informative first sentences, readers would have activated more concrete semantic representations of the target words (Hamada, 2013). The information would make it easier for them to integrate the meanings of the second sentences. On the other hand, readers experienced more difficulty achieving coherent comprehension from the first neutral sentence; therefore, the longer reading times in subsequent sentence would reflect their mental effort in inferring the target word meanings. This discussion would be to some degree a reliable one because, unlike the first sentence, the content of the second sentence was identical across conditions.

Finally, we will further discuss and summarize the reading behaviors of each of the two sentences in the opaque conditions. In the first sentences, they clearly noticed the semantic

inconsistency irrespective of context informativeness, as suggested by the longer reading times for the opaque conditions. However, though they might try to infer the meaning, it would be safe to conclude that they had not resolved the semantic inconsistency at this moment due to (a) the similar reading pattern irrespective of transparency, and (b) their continuing effort in the subsequent sentences, suggested by the longer reading times for the second sentences in the opaque conditions.

In this regard, it was possible while reading the first sentence that they suspended their judgment to collect more information in the subsequent sentence. The participants were notified in advance that they would read two-sentence passages; as a result, they knew they needed more context to infer the meanings. This is consistent with Huckin and Bloch's (1993) account of search for meaning processes, and has also been observed in a lexical inferencing study (Bengeleil & Paribakht, 2004) in which readers tried to use a wider context when they perceived that they could not find sufficient clues or difficulties in the immediate context. Therefore, most of the inflated reading times in the first sentence were possibly due to the reaction to the inconsistency between the two sources.

On the other hand, while reading the second sentence they would be trying to infer the actual or context-based meanings of the target words, and they made the mental effort according to the context quality of the first sentences. The next section examines how these reading behaviors would lead to their interpretations of opaque words.

What kind of semantic representations do intermediate-level Japanese EFL readers construct for unknown words through reading? (RQ2-2)

The analysis of the interpretation type produced for the comprehension questions in the reading task showed that readers made appropriate interpretations in most cases for transparent

words (Informative: 98.4%; Neutral: 96.4%), while the proportions were much lower for opaque words (Informative: 42.3%; Neutral: 33.0%).

However, the large proportions of None (Informative: 26.9%; Neutral: 25.0%) suggests the possibility that they were aware of the semantic inconsistency and avoided literal interpretations. The result that None was hardly found in the transparent conditions (Informative: 1.5%; Neutral: 2.8%) means that they were able to identify the possible answers to the comprehension question, at least in English. However, the semantic conflicts made it difficult for them to infer the words' actual meanings. Accordingly, the choice of None would be the result of readers' awareness of the semantic inconsistency and their intentional avoidance of literal interpretations. However, the above results seem to contradict the on-line reading time data, which suggested that they perceived the semantic inconsistency and also tried to infer contextual meanings while reading. This means that there was a large gap between attempts to infer context-based meanings and actually making appropriate interpretations.

There are some possible explanations of this discrepancy. The first is related to the lack of contextual support. In the present materials, the cumulative meanings derived from the two sentences did not narrow down to one specific meaning for the word even in the informative conditions. As a result, even if they tried to infer appropriate meanings, they could not come up with concrete semantic representations of the target words. Therefore, the insufficient clues might have led to inappropriate interpretations or to no response.

The importance of the contextual information was partly supported by the interpretation trend based on the context informativeness of the first sentence. When the first sentence contained informative clues about the target word meanings, the participants made less MBI (Informative: 29.2%; Neutral: 40.9%) and more CBI (Informative: 28.7%; Neutral: 19.5%). This means the quality of the initial encounter has a positive influence on subsequent interpretation. Concrete semantics derived from an informative first sentence would have made

it easier for them to integrate the content of the second sentence, which is consistent with the shorter reading times of second sentences in the opaque condition when the first sentence was informative context.

Interestingly, unlike Experiment 1, where PMI was more often made in neutral contexts (Informative: 22.1%; Neutral: 35.0%), this difference in proportions was not observed in this experiment (Informative: 13.7%; Neutral: 13.5%). This difference in the lower reliance on morpheme-based information for appropriate interpretations in this experiment might be attributed to the increased contextual information. In Experiment 1, limited information inferred from the single neutral context made them rely more on morpheme-based information. However, the addition of subsequent directive context would have reduced the effects of the context quality of the first sentence; as a result, the proportion of PMI was not affected. This is consistent with the above-mentioned result that the proportions of MBI differed by context in this experiment, but did not in Experiment 1 (Informative: 48.3%; Neutral: 46.7%). These results would reflect the readers' tendency to use morpheme-based meanings less for interpretations as available context information increased.

Because of these important roles of contextual support, their performance would have been better if additional and more directive contexts had been available. However, as noted in 3.2.1, the lack of sufficient clues is not a problem specific to this experiment, but rather of naturalistic text as a whole (Beck et al., 1983; Bensoussan & Laufer, 1984), and such remotely located clues are less likely to be used in inferences (Wesche & Paribakht, 2010). Therefore, this would not substantially affect the generalization problem. In addition, the possibility of context quality alone might not fully explain the smaller proportion of appropriate interpretations. This is because we confirmed in the norming study that readers could derive somewhat context-based meanings for opaque words.

Another possibility is the motivational importance of the target words in the task. In the experiment, they were asked to read the passages for comprehension. As a result, their attention was less likely to be directed to unveiling the meanings of unknown words. In the present experiment, the readers' goal was to provide written answers to the comprehension questions presented *after* reading. Therefore, the importance of the target word meanings they perceived *during* reading might have been relatively low. Accordingly, even if they tried to infer the meanings of the target words, their efforts would have been insufficient to come up with concrete meanings. If these questions had been given *before* reading, they might have made a greater effort to infer the targets. Nevertheless, as the focus of the present experiment was to examine the semantic memory representation of unknown words as a part of reading comprehension, this did not matter.

Although both factors (i.e., lack of contextual support, motivation) would account for the smaller proportions of appropriate interpretations, the latter factor would have had a greater effect, for the following reasons. First, the present materials were considered very easy for them in terms of vocabulary levels and syntactic complexity. Therefore, their constructed passage representations would not differ greatly, which cannot explain the performance differences among individuals (M = 37.69%, SD = 24.15, Min = 0.00%, Max = 87.50%).

Secondly, the great variation among participants in the proportion of appropriate interpretations would suggest their variability in perception of reading, including the way they should deal with unknown words. That is, some participants would have thought they should infer the meanings of unknown words, while others might have made less effort or ignored them (Fraser, 1999; Paribakht & Wesche, 1999). Although they were all asked to read for comprehension, how they actually perceived the instruction might have differed among individuals.

In sum, readers would not make sufficient mental effort to resolve semantic inconsistency between morphemes and context of opaque words, possibly because of the low importance of unknown words during reading. In addition, considering the low availability of clues beyond sentence boundary for inferencing, it can be argued that inappropriate interpretations for opaque words is less likely to be revised as readers proceed through a text.

3.2.5 Conclusion of Experiment 2

The purpose of Experiment 2 was to examine the semantic representations of unknown words after reading and their relationship with on-line text processing. To this end, intermediate-level Japanese EFL learners read two-sentence passages containing either transparent or opaque unknown words in the first sentences, and their interpretation of the target word was examined via a comprehension question immediately after reading.

The on-line reading time data showed that EFL readers were sensitive to the semantic relationship between morphemes and context, and possibly tried to infer the word meaning even after sentences containing unknown words by employing varying degrees of mental effort by the context quality of the preceding sentence (RQ2-1). However, this attempt did not always lead to contextually appropriate interpretations. They often interpreted unknown opaque words literally or did not produce any answers (RQ2-2). These results suggest that generating contextual meanings of opaque words while reading is uncommon for EFL readers.

The most plausible cause of this was a lack of motivation to interpret unknown words during the task. In normal reading, readers do not necessarily infer every unknown word unless doing so is critical for their task completion. However, at the same time, the variations in their performance would suggest that they set different standards for reading, and ways to deal with unknown words. This possibility of the effects of affective engagement was tested in Experiment 3.

3.3 Experiment 3: Use of Morphological and Contextual Information in Lexical Inferencing

3.3.1 Purposes and research questions of Experiment 3

Experiments 1 and 2 consistently showed that the difficulty of interpreting opaque words lies in generating context-based meanings going beyond the literal interpretations, even for intermediate-level learners. The participants often made morpheme-based interpretations even when the meanings were contradictory to surrounding contexts.

This low performance could be attributed to the lack of motivation for revealing the meanings of the target words while reading. Nonetheless, it remains unclear whether intermediate-level EFL learners could infer appropriately even when they were explicitly required to make lexical inferences. Although a past lexical inferencing study showed that learners at this level could effectively use both morphological and contextual information (Hamada, 2014), that study employed a multiple-choice format, which might have made the meaning generation process relatively easy. Therefore, the first purpose of Experiment 3 is to examine how well intermediate-level EFL learners infer unknown words in a lexical inferencing task instead of normal reading.

The second purpose was to reveal what makes the lexical inferencing of opaque words successful or unsuccessful. Although motivational factors could affect whether and how deeply readers try to process unknown words (e.g., ignoring, inferencing), it is the cognitive processes while inferencing that directly affect the outcomes of these inferences. Revealing the cognitive processes involved in accurately understanding unknown words, especially opaque words, will inform us how they should deal with those words. Since lexical inferencing is a subcomponent of general reading strategy (Wesche & Paribakht, 2010), the findings would be applicable to unknown word processing in normal reading.

As for measuring on-line cognitive processes, the measurement of reading times used in Experiment 1 and 2 was instrumental in determining whether or not they perceived the semantic relationship and also tried to infer the meaning (Brusnighan & Folk, 2012; Oakhill et al., 2016). However, it was not suitable for examining how they infer word meanings. Thus, Experiment 3 adopted the think-aloud method. Think-aloud comments produced while inferencing would reflect the kinds of linguistic clues (e.g., morphemes, context) and metacognitive behaviors (e.g., evaluating) they used (e.g., Hu & Nassaji, 2014; Huckin & Bloch, 1993; Wesche & Paribakht, 2010). In addition, qualitative analysis could elucidate why certain strategies were employed in a given situation.

Past L2 lexical inferencing studies have pointed out some useful strategies (e.g., confirming the inferred meanings; Hu & Nasaji, 2014), as well as problematic inference behaviors in general (e.g., overreliance on morphemes; Nassaji, 2003). It is generally agreed that effective strategy use is not the result of the use of one certain strategy but rather a combination of multiple strategies (Nassaji, 2003, 2006) and the use of strategies appropriate to a given situation (Bengeleil & Paribakht, 2004; Hu & Nassaji, 2014). In this regard, however, researchers have not yet fully examined what kinds of inferential strategies are effective in inferring opaque unknown words, which makes lexical inference problematic for many L2 learners (Huckin & Coady, 1999; Laufer, 1989). These findings hold important insights for Study 2 of this dissertation, which will target beginner-level EFL learners. Since they are considered to have more difficulty than immediate-level learners, revealing effective inferential strategies will tell us how to instruct them to infer unknown words.

The third objective is to examine why readers wrongly interpret unknown opaque words based on morphemes. Though this has not been empirically examined, past studies discussed it as arising from readers' lack of metacognition of their monitoring of the two sources: they decomposed unknown words into morphemes and finished inferencing without considering

context meaning (Bensoussan & Laufer, 1984; Hamada, 2014; Huckin & Bloch, 1993; Laufer, 1989). However, the participants in Experiments 1 and 2 were sensitive to the transparency/opacity of morphemes but often gave a literal interpretation. This raises the possibility that such inferences would be a result of how readers generate word meanings. Identifying the causes of such inferences would be helpful for teachers to help students when they made incorrect word-based inferences.

In sum, Experiment 3 examines (a) how well EFL learners infer unknown words when they are motivated to infer them, (b) inferential strategies that contribute to appropriate interpretations of opaque words, and (c) why some learners make word-based inferences even when they are semantically inconsistent with the surrounding context. These aims are summarized in the following research questions.

RQ3-1: How well do Japanese EFL learners infer unknown words in a lexical inferencing task according to the semantic transparency of the morphemes?

RQ3-2: What kinds of inferential strategies are related to the appropriate interpretations for opaque unknown words?

RQ3-3: What are the causes of inappropriate word-based interpretations of opaque words?

3.3.2 Method

3.3.2.1 Participants

The participants were 12 Japanese undergraduate and graduate students with various majors. They had learned English for at least six years in Japanese educational settings. Their estimated English proficiency was approximately at intermediate or more advanced levels based on their self-reported scores on English proficiency tests, such as the TOEIC Listening and Reading test (n = 3, M = 731.67, SD = 98.52, range = 630–865), TOEFL ITP test (n = 3, M

= 519.00, *SD* = 34.19, range = 490–567), and EIKEN test (pre-second grade: *n* = 2; second grade, *n* = 5). Their reading proficiency was tested with the reading proficiency test. The average score was 13.75 (*SD* = 2.92), ranging from 11 to 20 (Cronbach's α = .77).

3.3.2.2 Materials

The materials used in this study were identical with the 32 target compounds and paired single-context sentences used in Experiment 1 (Appendix 3). Again, the target words were pseudo compound target words (e.g., *pricewar*) whose surrounding contexts, either informative or neutral, were semantically consistent or inconsistent with the literal interpretations of the morphemes of the target words. Four presentation sets, in which every target word appeared in any of the four conditions, were prepared so that the participants would encounter eight target words from each condition. Unlike Experiments 1 and 2, filler sentences were not used because the purpose of this experiment was to investigate how participants processed and interpreted transparent or opaque unknown words when they were strongly motivated to infer their meanings, and there was thus no need to distract their attention.

3.3.2.3 Procedure

The experiment was conducted individually in a silent room. First, the author explained the general purpose of the study and outlined the entire session, and then the participants were randomly assigned to one of the four presentation sets.

Next, the participants took the lexical inferencing task. In this task, they were asked to report everything that they thought and did while inferring the meaning of the target word (boldface) embedded in the experimental sentences. To capture the participants' thoughts accurately, the experimenter did not interrupt their procedure unless they were silent for a

certain period of time (about ten seconds). In order to avoid the effects of spoken proficiency in English, these reports were made in their L1 (Japanese).

From the results of a pilot study (N = 2), it was found that (a) performing the dual tasks of speaking while inferring sometimes imposed cognitive burdens, and (b) sometimes only the inference outcomes were reported, or they tried not to report processes found irrelevant to arriving at their answers. Therefore, if they felt it difficult to report and infer simultaneously, the participants were allowed to engage in their inferential procedure for a short period of time (about 5 seconds) and then report everything they thought about during that silence. In addition, they were encouraged to report not only how they arrived at the meanings but also their thought processes that they perceived useless for inferring.

Even if they did not come up with concrete meanings for target words, they were encouraged to infer as much as they could. They were asked to include the inference outcomes and the meaning of the target words, but they were still allowed to change their answers until they moved on to next trials if they were not satisfied with their answers.

Each inference trial began with a "Ready?" screen on the PC. Then, when participants pressed a button on the keyboard, a context sentence appeared and they inferred the meanings of the target words while verbalizing their thoughts. When they finished inferring the meaning of a word, they pressed a button to move on to another trial. To familiarize the participants with this procedure, they performed a practice session for three target words. The target-context pairs for this session were adopted from previous research (Hamada, 2014; Webb, 2007). After that, the participants proceeded to the experimental session, and engaged in 32 trials while their voices were recorded. The presentation order of the target words was randomized. They were allowed to take a rest whenever they needed it. The think-aloud procedure lasted approximately 20 to 45 minutes. After the task, they were informed that none of the target words were real words. Finally, they took the reading proficiency test for 20 minutes.

3.3.2.4 Scoring and analysis

Outcomes of Lexical Inferences

Categorization of the inference outcomes (words meanings the participants responded while the inference task) was the same one with prior experiments (Appendix 2). Inference answers were classified into, *Morphology-based interpretation* (MBI), *Partially-morphology-based interpretation* (PMI), *Context-based interpretation* (CBI), *Inappropriate interpretations*, and *None*. MBI, PMI, and CBI were considered semantically appropriate for transparent words, while PMI and CBI were considered appropriate for opaque words. The author and a graduate student majoring in English language education first independently scored 30% of the data, resulting in 92.2% agreement. After discrepancies were resolved through discussion, the remaining data was scored by the author.

Inferential Strategy Use

First, the recorded protocols were transcribed. Due to some recording and presentation errors, 12 attempts (3.13%) were removed from the analysis. Next, an inferential strategy list was developed using an inductive procedure; all the protocols were reviewed, and identified strategies used in past studies (Hu & Nassaji, 2014; Nassaji, 2006) were included in the list, resulting in seven strategies: Analyzing, Partial analysis, Paraphrasing, Inferencing from context, Confirming, Questioning, and Stating difficulty or failure. These strategies were further broadly categorized into four types: (a) form-focused strategies, (b) meaning-focused strategies, (c) evaluation, and (d) monitoring based on the nature of the strategies (Hu & Nassaji, 2014; Nassaji, 2006). While form-focused and meaning-focused strategies are used to identify the meanings of unknown words, evaluating and monitoring strategies are related to metacognitive strategies that control or monitor their inferential processes. Definitions and examples of the strategies are shown in Table 3.16.

Table 3.16

Definitions and Examples of Inferential Strategies

1. Form-focused Strategies

(a) Analyzing

Definition: Analyzing a word according to the literal meanings of its two morphemes

Example: ***sealine***… *sea* means sea and *line* means line, so it might be a coastline.

(b) Partial analysis

Definition: Relying on the literal meaning of one of the pseudo compound word's morphemes or trying to figure out metaphoric meanings for the morphemes

Example: ***waterwear***…*water* is water, and water is cold…so it's like Cool-biz. / *he gave a* ***trainstop*** *as a birthday present*…Uh, probably he gave a toy train.

2. Meaning-focused Strategies

(a) Paraphrasing

Definition: Paraphrasing or translating the part of the text that contains the target words.

Example: This part means they decided to hold a meeting during something.

(b) Inferring from context

Definition: Guessing the meaning of the target words by using the surrounding context clues.

Example: This word may be related to a kind of job because it says "want to be a."

3. Evaluating Strategy

(a) Confirming

Definition: Examining the appropriateness of the inferred meanings by using the information in the text.

Example: …it must be appropriate as it converges with what the sentence says.

(b) Questioning

Definition: Questioning their own inferences.

Example: I think it is related to something enjoyable, umm… is it right?

4. Monitoring Strategies

(a) Stating difficulty or failure

Definition: Making statements about the failure of inferencing or the difficulty of the target words

Example: Umm, it is difficult to infer, I have no idea what it means. / It doesn't make sense at all, it must be wrong.

Referring to the list, two raters independently categorized 30% of the protocols, with 87.1% inter-rater agreement; discrepancies were resolved through discussion and the author categorized the remaining data.

Form-focused strategies were those related to the use of the target word's morphological information. Analyzing (analyzing a word according to the literal meanings of its two morphemes) and Partial analysis (relying on the literal meaning of one of the pseudo compound word's morphemes or trying to figure out metaphorical meanings for the morphemes) belonged to this category. It should be noted that the category of Partial analysis was created for this study. In other studies (Hu & Nassaji, 2014; Nassaji, 2006), partial use of the morphemes of unknown words (e.g., prefix, suffixes) was also classified as Analyzing; however, partial use of morphemes in the present materials would reflect the participants' attempts to derive contextually appropriate interpretations for opaque words, which are also reflected in PMIs. For this reason, this study added the Partial analysis strategy to reflect learners' different cognitive processes.

Meaning-focused strategies entail use of semantic information of the context apart from the target words. Paraphrasing (paraphrasing or translating the part of the text that contains the target words) and Inferring from context (guessing the meaning of the target words by using the surrounding context clues) were found.

Evaluating strategies were used to evaluate the appropriateness of the inferences. Two strategies, Confirming (examining the appropriateness of the inferred meanings by using the information in the text) and Questioning (questioning their own inferences), were included.

Finally, Stating difficulty or failure (making statements about the failure of inferencing or the difficulty of the target word) belonged to Monitoring strategies, and was related to their own inferential process.

After coding the protocols, each inferential strategy's probability of being used was calculated. This was done by calculating the probability of whether a certain strategy was employed per inference trial for each condition. To compare inferential strategy use by semantic transparency and context, a non-parametric Wilcoxon signed-rank test was performed; the interaction of the two variables was not considered due to the small sample size. Given the effects of multiple comparison, the alpha level was adjusted to .0125 with Bonferroni adjustment (.05/4).

3.3.3 Results

3.3.3.1 Interpretations of target words

First, the proportions of each interpretation type are presented in Table 3.17. There was a clear interpretation pattern difference according to the transparency of morphemes. The participants predominantly made MBIs for transparent words (Informative: 96.4%; Neutral context: 84.4%) and inappropriate interpretations were hardly seen (Informative: 0.0%; Neutral: 1.0%).

Table 3.17

Proportions of Each Interpretation Type (%) in the Lexical Inference Task in Experiment 3

	Informative context					Neutral context				
Transparency	MBI	PMI	CBI	INI	None	MBI	PMI	CBI	INI	None
Transparent	96.43	3.57	0.00	0.00	0.00	84.38	10.42	4.17	0.00	1.04
	(9.24)	(9.24)	(0.00)	(0.00)	(0.00)	(16.85)	(15.63)	(8.43)	(0.00)	(3.77)
Opaque	7.29	37.58	49.65	2.90	2.31	11.46	60.71	13.84	7.76	6.23
	(11.56)	(22.08)	(24.92)	(8.21)	(6.16)	(8.76)	(21.90)	(17.76)	(10.25)	(8.09)

Note. Standard deviations are in parentheses.

On the other hand, the interpretation distribution was not uniform for opaque words (see Figure 3.8); that is, PMIs and CBIs, contextually appropriate interpretations, accounted for most cases (Informative: 87.4%; Neutral context: 74.6%), and MBIs were observed about 10% of the time (Informative: 7.3%; Neutral context: 11.5%). As for the effects of context, descriptively speaking, CBIs were more often produced for informative contexts (Informative: 49.7%; Neutral: 13.9%) and PMIs for neutral contexts (Informative: 37.6%; Neutral: 60.7%); this trend replicated the results of Experiment 1 that used the same single sentence frame. Overall, the participants made appropriate interpretations for not only transparent words but also opaque target words.

3.3.3.2 The inferential process and its relationship with outcomes

This section focuses on the lexical inferencing strategy use and explores its relationship with the outcomes of lexical inferencing. Originally this analysis was intended to compare the inferential strategy use of appropriate and inappropriate inferences. However, due to the small proportion of inappropriate interpretations for opaque words (Informative: 5.2%; Neutral: 14.0%), we instead compared the inferential strategies that the participants employed for transparent and opaque targets in general. If it was observed that some strategies were used more often based on the semantic transparency of the target words, the strategy is considered relevant to the appropriate interpretation.

Descriptive statistics of the inferential strategies used are presented in Table 3.18 (see also Figure 3.7). In general, strategies used to generate meanings (Form-focused, Meaning-focused strategies) were employed more often than metacognitive strategies (Evaluating, Monitoring strategies).

Table 3.18

Descriptive Statistics of Inferential Strategy Use (%) in the Lexical Inference Task in Experiment 3

		Informative context			Neutral context		
Strategy	Transparency	*M*	95% CI	*SD*	*M*	95% CI	*SD*
Form-focused							
Analyzing	Transparent	96.73	[92.21, 100.50]	5.69	90.63	[83.66, 97.59]	10.36
	Opaque	60.07	[45.87, 74.27]	21.14	59.67	[43.46, 75.89]	24.14
Partial analysis	Transparent	5.51	[−0.07, 11.08]	8.29	19.79	[7.21, 32.37]	18.72
	Opaque	50.69	[35.41, 65.98]	22.75	70.09	[53.48, 86.70]	24.72
Meaning-focused							
Paraphrasing	Transparent	73.92	[61.60, 86.23]	24.93	75.83	[63.62, 88.05]	28.16
	Opaque	86.46	[74.85, 98.06]	16.54	75.60	[60.65, 90.54]	22.25
Inferencing from	Transparent	25.74	[12.63, 38.85]	19.51	21.92	[7.20, 36.65]	21.92
context	Opaque	68.75	[57.64, 79.86]	16.54	54.02	[37.95, 70.08]	23.91
Evaluating							
Confirming	Transparent	19.05	[7.79, 30.30]	16.76	18.75	[10.01, 27.49]	13.01.
	Opaque	6.25	[0.83, 11.67]	8.07	14.58	[5.62, 23.55]	13.34
Stating difficulty	Transparent	3.57	[−4.39, 11.53]	11.85	5.21	[0.17, 10.58]	8.00
or failure	Opaque	24.65	[10.22, 39.08]	21.28	20.09	[10.63, 29.55]	18.57
Monitoring							
Questioning	Transparent	6.55	[−1.09, 14.19]	11.37	3.13	[−0.51, 6.76]	5.41
	Opaque	11.46	[0.91, 22.00]	15.69	20.98	[8.50, 33.46]	18.57

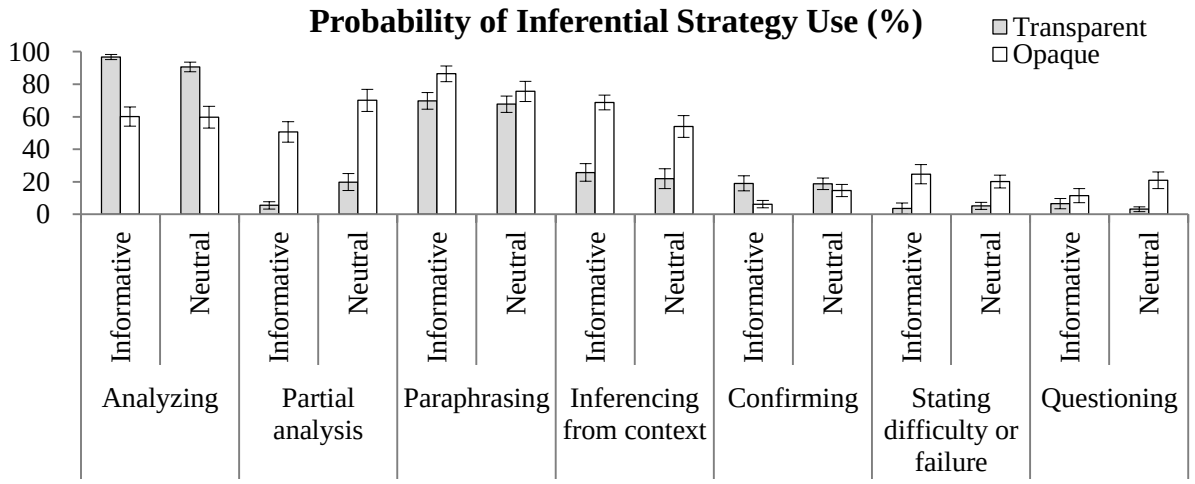


Figure 3.7. Probability of inferential strategy use as a function of semantic transparency and context informativeness in Experiment 3.

To compare the difference in strategy use for transparent and opaque target words, a non-parametric Wilcoxon Signed-Rank Test was performed on the probability of strategy use by semantic transparency and context. The results revealed that five (out of a total of seven) strategies were used differently as a function of transparency and context (see Table 3.19).

Table 3.19

Summary Table for Wilcoxon Signed-Rank of the Effects of Transparency and Context on the Proportions of Inferential Strategy Use in Experiment 3

Strategy	Factor	Type of difference	*Z*	*p*	*r*
Analyzing	Transparency	Transparent > Opaque	3.06	.002	.63
	Context	None	0.54	.593	.11
Partial analysis	Transparency	Transparent < Opaque	3.06	.002	.63
	Context	Informative < Neutral	2.94	.003	.60
Paraphrasing	Transparency	None	1.97	.049	.40
	Context	None	1.48	.138	.30
Inferencing	Transparency	Transparent < Opaque	3.49	.000	.71
from context	Context	None	1.87	.062	.38
Confirming	Transparency	None	1.97	.049	.40
	Context	None	1.28	.201	.26
Stating failure	Transparency	Transparent < Opaque	2.67	.008	.55
or difficulty	Context	None	0.31	.758	.06
Questioning	Transparency	Transparent < Opaque	2.55	.011	.52
	Context	None	1.59	.112	.33

Below, we will review each of them and its relationship with inference outcomes, and see how these strategies were used with reference to some exemplified inferential protocols (translated into English by the author). First, Analyzing (i.e., analyzing a word according to the literal meanings of its two morphemes) was used in almost all cases for transparent targets irrespective of context, which was more frequent than for opaque targets ($Z = 3.06$, $p = .002$, r

= 63). This supports the results of the inferencing outcomes trial in that participants made more MBIs for transparent words as well.

However, note that although this strategy was used about 60% of the time for opaque targets (Informative: 60.1%; Neutral: 70.1%), that would not necessarily mean that participants did not morphologically analyze targets the rest of the time, given the automatic nature of this process (Fraser, 1999; Pollatsek et al., 2008, 2011). Let us examine the following protocol produced for opaque target.

> Example 1
>
> …***housewear*** on that day, umm, no one did not approach ***housewear*** because a large shark appeared….***housewear***, umm, then the word means a beach because it must be the place a shark appears. (excerpt produced for *There was a large shark so nobody went near the* ***housewear*** *on that day.* [opaque/informative])

The participant processed the target word repeatedly but did not decompose it into morphemes (i.e., *house, wear*). However, it appears that she avoided analyzing it, possibly because she felt that the literal meaning would be unrelated to the actual meaning due to the semantic conflicts of the two sources, and decided not to report the analytic process, whether consciously or unconsciously.

Second, Partial analysis (i.e., relying on the literal meaning of one of the pseudo compound word's morphemes or trying to figure out metaphoric meanings for the morphemes) was used more often for opaque words than for transparent ones ($Z = 3.06$, $p = .002$, $r = .63$). Because it was impossible for the participants to achieve a meaningful interpretation through a literal understanding, participants tried to rely on a part of the pseudo compound word's morphemes or to generate metaphorical meanings for the morphemes that fit in the context.

Example 2

I do not know what the shop closed because of *businessbag* means…business and bag, umm … a hard bag for business. What is this, why? … We say we are in blind alley when we are in trouble and cannot do anything. So it is something like a helpless situation, I think. (excerpt produced for *The shop closed after the hard* ***businessbag*** *in the area*. [opaque/informative])

In Example 2, the participant first noticed with some confusion that the morpheme-based meaning was not meaningful in the context. Then he resolved it by using the associative meaning of one morpheme that is used in Japanese (bag → blind alley [*fukuro koji* in Japanese]), which is also consistent with the contextual meaning. Therefore, Partial analysis was used to achieve a meaningful interpretation of context while retaining part of the morpheme-based meanings.

Furthermore, this was the only strategy that showed a different trend depending on Context: it was used more often for neutral contexts than for informative contexts ($Z = 2.94$, $p = .003$, $r = .60$). The lack of sufficient information in neutral contexts would have led the participants to rely on morphological information, even if it was partial. This would also be consistent with the results that PMIs were more often made in neutral contexts (60.7%) than in informative contexts (37.6%).

As for Meaning-focused strategies, Inferring from context (i.e., guessing the meaning of the target words using the surrounding context clues) was more frequently used for opaque targets than for transparent ones ($Z = 2.67$, $p = .008$, $r = .55$). Let us consider the following example).

Example 3

Sara talked with a *gunball*. …*gunball*, *gunball*, gun and ball… a bullet? It cannot be that. … She talked with a …it probably means a person because it refers to a subject she talked on the phone. And a person we talked at night is usually a friend, or someone like that. So, the meaning would be a friend. (excerpt produced for *Sara talked with a* ***gunball*** *on that night.* [opaque/neutral])

As in Example 2, the participant first noticed that the literal interpretation was strange within the context. Then she rejected it and later made an appropriate interpretation. The fact that morphology-based meanings did not fit the context prompted her to narrow down the possible meaning of target words based on the context. On the other hand, for transparent words, there was no need to generate or narrow down the meanings of unknown words because the two sources semantically converge; thus, they only had to choose morpheme-based meanings for transparent words.

However, no significant difference was seen for the other Meaning-focused strategy, Paraphrasing (i.e., paraphrasing or translating the part of the text that contains the target words). Although this strategy entailed semantic processing of context, simply paraphrasing the context (i.e., translating English into Japanese) would not contribute to resolving the semantic inconsistency between the two sources. This account can be supported by the high frequency of this strategy use overall irrespective of conditions (73.9–86.5%).

With regard to metacognitive strategies (i.e., Monitoring, Evaluating), the participants made more comments in the categories of Stating difficulty or failure (i.e., making statements about the failure of inferencing or the difficulty of the target words) and Questioning (i.e., questioning their own inferences) for opaque targets than for transparent targets (Stating difficulty or failure: $Z = 2.67$, $p = .008$, $r = .55$; Questioning: $Z = 2.55$, $p = .011$, $r = .52$).

Examples of these strategies can be seen in the previous example protocols produced for opaque targets when they were faced with semantic conflicts of the two sources. Below is another example where these strategies are used in a different way.

> Example 4
>
> It sounds strange that one lost her home because of a book… it is the cause of losing one's house. Umm, so it might be natural disaster, such as an earthquake. Umm, design, and book, I'm not sure it really is an earthquake, right? It must not be a sketchbook, but it is very difficult. But I think it means an earthquake. (excerpt produced *Sara lost her home because of the big* ***designbook*** *five years ago.* [opaque / neutral])

In Example 4, the participant was able to make a contextually appropriate interpretation (earthquake) by using contextual information. However, he was still uncertain the meaning was correct and was skeptical about the correctness of the inference. Thus, semantic conflicts of opaque words would prompt participants to monitor their inferential process even after making inferences.

However, Confirming (i.e., examining the appropriateness of the inferred meanings by using the information in the text), also a metacognitive strategy, did not differ by Transparency (Z = 1.97, p = .049, r = .40). Rather, it was used slightly more often for transparent words (Informative: M_{diff} = 12.8%; Neutral: M_{diff} = 3.9%). Originally, it was assumed that the difficulty of inferencing opaque words might increase the necessity of evaluating the appropriateness of inferred word meanings. One possible explanation is related to the linguistic sources available for evaluation. Let us consider the following example with a transparent target word.

Example 5

The sport shop sold a *waterwear*, …during summer sale at a low price. The meaning is a swimsuit I think. *Water* is water and *wear* is something to wear. And it must be right because the word is related to what the sport shop sells. (excerpt produced for *During the summer sale, the sport shop sold a* ***waterwear*** *at a low price*. [transparent /informative])

In Example 5, the participant first browsed the whole context, and then interpreted the target word literally. Then he confirmed the correctness of his inference by considering the contextual meanings again. In this way, the semantic consistency of transparent words enabled participants to evaluate meanings inferred either from morphemes or from context using the other source.

On the other hand, in case of opaque words, the participants had to use both sources of information in the process of inferencing appropriately; as a result, it might have been less likely for them to use this same information for further evaluation (Huckin & Bloch, 1993). Indeed, they considered both morphological and contextual meanings before making inferences and finished their inferential processes quickly, as in Examples 2, 3, and 4. Therefore, the ease of evaluation for transparent words might have overridden the necessity of using this strategy for opaque words.

In sum, it was found that readers changed their inferencing processes depending on the semantic relationship between morphemes and context. This flexible strategy use led to their overall good performance on inference outcomes.

3.3.3.3 Qualitative analysis of morpheme-based interpretations for opaque words

The analysis aimed at exploring the possible causes of the most typical mistakes in inferencing—that is, making word-based inferences (MBIs) for opaque words. As noted before,

the number of MBIs for opaque words was quite small: 16 cases out of 188 attempts (8.5%). Therefore, this analysis adopted a qualitative procedure: all the MBIs were reviewed and possible causes identified. In classification, one criterion, whether the participants were aware of the inconsistency between the two sources of information, was applied because it has been thought to be a possible cause of literal interpretation (Hamada, 2014). As a result, the following five types of causes were identified: (a) Negligence of context, (b) Lack of evaluation, (c) Persistent reliance on morphemes, (d) Modification of context, and (e) Elaboration of context. Table 3.20 provides the definition of each category. Participants made the first two types of mistakes without awareness of the semantic inconsistency, but for the other three participants exhibited difficulties induced by the inconsistencies. Next, we will examine all the categories and see how they resulted in such inferences, analyzing some actual think-aloud protocols.

Table 3.20

Categories and Definitions of Causes of Literal Interpretation for Opaque Words in Experiment 3

Category	Definition
(a) Negligence of context (*n* = 4)	Readers made a literal interpretation of morphological information without considering contextual meanings.
(b) Lack of evaluation (*n* = 2)	Readers used the literal meanings of morphemes in interpreting target words and failed to notice the semantic inconsistency conveyed by the morphemes and context.
(c) Persistent reliance on morphemes (*n* = 4)	Readers noticed the semantic inconsistencies but persisted in using the literal information of morphemes, sometimes due to the difficulty involved in deriving contextually appropriate meanings.
(d) Modification of context (*n* = 3)	Readers changed some of the interpretation of the surrounding context so that the literal meanings of the morphemes made sense in the context.
(e) Elaboration of context (*n* = 3)	Readers made additional interpretation of context so that the literal meanings of morphemes made sense in the context.

The first category, *Negligence of context*, covers cases in which participants made inferences solely relying on morphological information (n = 4). In this case, participants used literal morphological analysis at some point in the inferencing process, and then finished it without referring to contextual information, as can be seen in the following example.

> Example 6
>
> … *don't know how to use a guardperson*. By translating it literally, it means a person who guards others I think, like a guard man. *Guard* means protect and *person* means people. So, it is like an SP, OK. (excerpt produced for *Susan didn't know how to use the* ***guardperson*** *when she was 12.* [opaque / neutral])

In Example 6, he read aloud the phrase containing the target word, and then interpreted it literally by decomposing it into morphemes, and finished inferring the meaning without considering contextual meanings. Thus, the cause was ignorance or overreliance on the meanings of the morphemes and failure to consider the context.

The second type of cause, *Lack of evaluation* (n = 2), is similar to the first one, but in this case, participants used the context to some degree. In this type of inference, participants translated all or part of the context including the target word, but failed to notice the semantic discrepancy between the context and morphemes.

> Example 7
>
> Umm, *summer sale*, during summer sale, the sport shop, at a low price, *sold a lunchhour*, it sold something at a low price, …it is lunch time I think because *lunch* means daytime and *hour* is time or period. (excerpt produced for *During the summer sale, the sport shop sold a* ***lunchhour*** *at a low price.* [opaque/informative])

In Example 7, The participant translated the phrase that contained the target word and interpreted the target literally. Since participants correctly understood what the context said and fit the literal meaning of the target word into it, the problem here was that they failed to evaluate the semantic relationship between morphological and contextual information even when they took both morphemes and context into account.

The third type of error, *Persistent reliance on morphemes* (*n* = 4), was possibly caused by inflexible adherence to morphological information. Participants were aware of the inconsistency between the target word's morphemes and its context; however, failure to infer possible meanings made them use morpheme-based meanings, sometimes reluctantly.

Example 8

guardperson, *guardperson*, don't know how to use… Umm, what is *guardperson*? It is unlikely a girl doesn't know how to use a guard man. … But, it is like a guard man because it is guardperson, though I'm not sure. (excerpt produced for *Susan didn't know how to use the* ***guardperson*** *when she was 12.* [opaque/neutral])

In Example 8, the participant first thought the literal interpretation did not make sense. However, after a while she made the literal interpretation, though she was not confident in her answer. Thus, this kind of interpretation was caused even when the semantic inconsistency was perceived. Participants might have had difficulty suppressing the morphological information and prioritizing context-based meanings.

Unlike the above three types, the following two inferencing processes were different in that participants attained some kind of confidence in their literal, morpheme-based interpretations by changing the interpretation of the context. In the fourth type, *Modification of*

context (*n* = 3), participants tried to make literal interpretations meaningful by changing the interpretation of part of the context. Let us examine an example.

> Example 9
>
> *He did not use*… Umm, it might be an object as it says he did not use. Umm, it may be a native place as in a hometown. So, *did not use* might mean he did not visit I think. (excerpt produced for *Cameron had never used a* ***homecity*** *in his life.* [opaque/neutral])

In Example 9, she first narrowed down the word meaning from the context (i.e., an object). Then she decided on a literal interpretation by modifying a word meaning in the context (use → visit) that often collocates with the literal interpretation of the target (native place). As a result, she could achieve a meaningful representation of the sentence by changing the interpretation of the context.

Finally, *Elaboration of context* (*n* = 3) refers to cases in which participants intentionally selected a literal interpretation but did not change the contextual meaning. They tried to maintain the coherence of the two sources of information by making additional interpretations of context, exemplified in the protocol below.

> Example 10
>
> It is strange one did not feel the designbook, it is total nonsense. … but possibly, it means a designbook I think. Maybe, he had seen many designbooks before, … and when he saw a designbook that night, he felt that this was not the kind of designbook he had seen before. …Then, it's possible Shorn didn't feel the designbook. (excerpt produced for *Shorn didn't feel the* ***designbook*** *on that night.* [opaque/neutral])

In Example 10, after noticing the conflict between the morphemes and context, he adopted the literal interpretation by reasoning that Shorn would have had ample experience in seeing *designbooks*, and the one he saw on that night was not a typical *designbook*; as a result, the literal interpretation became plausible in the context for the participant. Therefore, he might have tried to resolve the semantic inconsistency by inferring an underlying scenario that allowed the conflicting morphological and contextual meanings to co-exist.

In sum, although some word-based inferences were caused by insufficient use of contextual information or lack of monitoring, as has been pointed out (Hamada, 2014; Huckin & Bloch, 1993; Nassaji, 2006), it is worthwhile to note that such inferences were made after the perception of the semantic inconsistency or as a result of readers' mental effort to make the literal interpretations meaningful by changing the interpretation of the context.

3.3.3.4 A follow-up analysis: Individual patterns of literal interpretations of opaque Words

However, given the smaller proportion of literal interpretations, there arises another question. That is, are these literal interpretations always a problem of their inferential skills? Of the five types, Negligence of context and Lack of evaluation were caused by the insufficient monitoring of the two sources. On the other hand, the other three can in a sense be seen as different responses to semantic inconsistencies because these interpretations were all made after the participants were aware of the semantic inconsistency. Therefore, it is worthwhile considering whether they could have made appropriate interpretations (i.e., PMI, CBI) if they had been asked to revise the literal interpretation without changing their interpretation of the context. Thus, we will explore individual performances, because if the problem lies in inferential skills, the same individual will make similar interpretations consistently throughout the task. This was done by calculating the number of literal interpretations for opaque words of

each individual, and the matter of when they were made was also taken into consideration in the 16 encounters of opaque targets. The results are shown in Appendix 5. The number of literal interpretations made per individual were quite small, ranging from 0 to 4 times. This means that in most cases they did not make interpretations overall relying exclusively on the literal interpretation of opaque words. It was also noticeable that most of these interpretations were made in the early part of the lexical inference task: 10 out of 16 (62.5%) were made by the fourth encounter, suggesting that participants did not know how to resolve the semantic inconsistencies at first, but learned to avoid such interpretations as their exposure to opaque words increased. This is consistent with the fact that Persistent reliance on morphemes (n = 4), whose cause was possibly the lack of a resolution of the inconsistency, were all made on either the first (n = 3) or second (n = 1) encounter with opaque targets. Accordingly, this smaller proportion of literal interpretations for opaque words in each participant and changing trend with exposure would indicate that these interpretations were not necessarily due to their inferential skills.

3.3.4 Discussion

How well do EFL learners infer unknown words in a lexical inferencing task according to the semantic transparency of the morphemes? (RQ3-1)

The first research question of Experiment 3 concerns lexical inference performance, especially for opaque target words. The results showed that in most cases the participants made appropriate interpretations of opaque words (Informative: 87.4%; Neutral context: 74.6%). In addition, they made hardly any literal interpretations (i.e., MBI), which were often seen in Experiment 1 with the same materials.

This relatively good performance in this experiment was consistent with the initial prediction and can be explained in light of the motivational importance of the target word

meaning for task completion. In Experiment 1, although translating sentences indirectly required inferring the meanings of unknown words in context to produce translations, their attention was more directed to the message of the entire sentence rather than to the meaning of the target word alone.

On the contrary, in this experiment, the sole goal of the lexical inferencing task was to accurately infer the meanings of target words. The participants had to generate the meaning of the target word for task completion. As a result, it was possible that the participants were more affectively engaged in the inferencing process and thus succeeded in generating contextually appropriate meanings. As can be seen in the think-aloud comments, the participants spent much time inferring the meanings by employing various inferential strategies. Overall, the participants in this study were able to infer the meanings of target words based on the semantic relation between morphological and contextual information. In addition, they tried to infer the meanings by focusing on the target word only, rather than changing their interpretations of surrounding context. This is in line with previous studies that insisted on the importance of the affective factor in whether and how deeply readers process unknown words (Laufer & Hulstijn, 2001; Hulstijn & Laufer, 2001; Paribakht & Wesche, 1999; Wesche & Paribakht, 2010).

What kind of inferential strategies are related to the appropriate interpretations for opaque unknown words? (RQ3-2)

In order to explore the cognitive processes that were peculiar to the appropriate interpretation through semantic transparency of unknown words, we compared the inferential strategy use as a function of transparency and context because of the overall high proportion of appropriate interpretations.

The results of Wilcoxon's signed rank test showed that a total of six strategies were used differently according to either semantic transparency or contextual informativeness.

Specifically, (a) Analyzing was more often used for transparent targets, (b) Partial analysis, Inferencing from context, Stating difficulty or failure, and Questioning were more often used for opaque targets, and (c) Partial analysis was more often used in Neutral contexts. Also, the analysis of think-aloud comments showed why such strategies were used differently.

The discussion here focuses on how readers should deal with opaque unknown words in terms of inferential strategy use. In so doing, we will explain the roles of the meaning generation process and metacognitive controls and their time course in inferencing (Huckin & Bloch, 1993).

In generating contextually appropriate meanings of opaque targets, the participants used part of morphological information and interpretations solely based on contextual meanings, and this is consistent with their inference outcomes (i.e., PMI, CBI). These strategies were used more frequently for opaque targets than transparent ones. This is quite natural given that the literal interpretations of the morphemes yielded semantically implausible word interpretations. Thus, their perception of the semantic inconsistency made them use strategies that did not rely on the literal meanings of morphemes.

In so doing, they used part of the morphological information that can be compatible with the sentence meanings (i.e., Partial analysis), or inferred it solely from context (i.e., Inferring from context). The two strategies both reflect readers' mental efforts to explore context-based meaning. It can be said that the use of these strategies directly contributed to the appropriate interpretations because the sources of information involved in the strategy are identical with PMIs and CBIs, respectively.

However, the outcomes of inference were partially inconsistent with the strategy use when context informativeness was considered. While a higher proportion of PMI in neutral contexts is supported by the more frequent use of Partial analysis for neutral contexts, that of Inferring from context did not differ by context despite the larger proportion of CBI in informative contexts.

This discrepancy might be related to the way this strategy is used. In Example 2, this strategy was used to generate the meanings of target words, which resulted in the final interpretation of the target words. However, the next protocol would indicate this strategy was used slightly differently.

> Example 11
>
> … the word would mean her job because it says *wanted to be a*, …if so it is strange that she wanted to be a carpet. Then, it would be a waiter because *floor* in the word mean some space, and I have an image that a waiter deals with space of a restaurant. (excerpt produced for *Julia wanted to be a* ***floorcover*** *when she was young.* [opaque/ neutral])

In this example, the participant first inferred that the word is related to her job from the meaning of the context (underlined), and then she felt it strange because the inferred meaning was inconsistent with the word-based interpretation of the target word. Then she arrived at an interpretation that included partially morphological information.

In this inferential process, she used the initially inferred meaning (i.e., job) to evaluate the consistency of the two sources, and then tried to explore the contextually appropriate meaning using the associations of part of the morphemes. Therefore, Inferring from context can be used not only for the final decision of the word meaning (as in Example 3), but also for initial, tentative inference. In addition, as in the above example, such tentative inferences can be made for neutral contexts, even though these allow a wide range of possible meanings. That is one reason that the frequency of Inferring from context did not differ by context informativeness, due to the multiple ways in which the strategy can be used. Although this kind of report was not always included in the protocols, it is reasonable to insist that perception of

the inconsistency involves readers' prediction of the context-based meaning, which is later validated with the morpheme-based meaning.

Now we will examine how metacognitive strategies were used when inferring opaque target words. The participants expressed difficulty and questioned their inferences for opaque words (i.e., Stating difficulty or failure, and Questioning) more often than for transparent ones. The analysis of the protocols showed that these strategies were made both while and after making inferences (Example 4). The primary source of their difficulty was the semantic conflict between the two sources. Although these strategies were used rarely, given the large proportion of appropriate interpretations, it would be readily expected that they might have experienced certain processing difficulties, whether at the conscious or unconscious level. The low frequency was partly due to the nature of the think-aloud method, in which cognitive processes that are difficult to verbalize were less likely to be observed.

Although the use of these strategies itself does not directly affect the outcome because they are not related to identifying the meanings of words, metacognitive awareness would guide readers to explore context-based meaning (Hu & Nassaji, 2014; Huckin & Bloch, 1993). After facing processing difficulties, participants continued to explore other interpretations of the word, going beyond literal interpretations. Such awareness would lead to the use of an effective strategy (Partial analysis, Inferring from context) instead of persisting in a literal interpretation.

The result that Confirming, a metacognitive strategy, was not used differently by semantic transparency was a little surprising. However, the analysis of inferential protocols (Example 2, 3) showed that the readers often used both morphological and contextual information in the meaning generation process; as a result, they were less likely to feel the necessity of evaluating their inferences. This is consistent with the behavior whereby readers try to use different sources of information when they feel processing difficulty (Bengeleil & Paribakht, 2004; Huckin & Bloch, 1993). According to these studies, a failure of inferences

from the immediate context prompts their metacognitive control process to search for more clues. Therefore, when they think their inferences are plausible in the use of the two sources for inference in the case of present study, they would not try to use the same information.

Moreover, this result does not contradict the suggestions about the importance of evaluation (Hu & Nassaji, 2014; Nassaji, 2003, 2006; Nation, 2013). They insist that evaluative processes are important because a lack of evaluation results in less accurate or incorrect inferences, as readers do not consider other useful sources. Accordingly, it can be argued that only if they consider the available sources, evaluation *after* inferencing is not necessarily a mandatory process. Thus, the meaning generation process, which requires the examination of the two sources *during* inferencing, would have played the part of an evaluative role.

To sum up, the appropriate interpretation of opaque words is related to readers' awareness of the semantic inconsistency, which would trigger them to explore the context-based meaning by integrating the two sources, or fully relying on the context-based meaning. However, the process is not always straightforward: they sometimes make tentative inferences and later refine or abandon them, or monitor their inferences both during and after inferencing.

What are the causes of inappropriate word-based interpretation for opaque words? (RQ3-3)

The analysis of the literal interpretation of morphemes for opaque words revealed that it has various causes (i.e., Negligence of context, Lack of monitoring, Persistent reliance on morphemes, Modification of context, Elaboration of context). Through an examination of inferential protocols, we analyzed how they led to these interpretations. These causes differ in whether or not they perceived the semantic inconsistency and, if so, how they behaved after this perception. We will discuss which cognitive processes led to these results with reference to the findings of past studies.

The first two causes (i.e., Negligence of context, Lack of monitoring) would be due to a lack of monitoring the surrounding context or its interrelationship with morphemes. These are the causes most frequently reported in past studies (e.g., Laufer, 1989; Nassaji, 2003, 2006). In the case of Negligence of context, participants' attention was drawn to word morphemes so strongly that they failed to consider the context-based meaning: they finished inferring after the morphological analysis. Furthermore, even when they considered both sources, the evaluative system sometimes did not function properly, as found in Lack of monitoring. This is consistent with studies that pointed to the important roles of evaluating processes (e.g., Hamada, 2014; Hu & Nassaji, 2014; Nassaji, 2006).

In the case of Persistent reliance on morphemes, the problem lies in how they behave after noticing the inconsistency. The participants were aware that the morpheme-based meaning did not make sense but had no idea how to resolve the problem. As a result, they end up deciding on a morpheme-based interpretation. Such inflexibility of interpretations was also found in previous studies in which readers persisted in using a known word meaning even when the word is used with a different, unfamiliar meaning (Ushiro et al., 2010). These studies have discussed the fact that changing an interpretation based on the context requires flexibility in changing interpretations and a certain level of proficiency.

Modification of context might explain why distortions were found in Experiment 1: Participants distorted the context meaning so that the word-based interpretation made sense. As seen in Example 9, readers made such a decision by noticing the inconsistency and then trying to change the interpretation of context rather than the literal meanings of the target words. Similar findings can be seen in a previous study (Bensoussan & Laufer, 1984) that reported that unconscious incorrect inferences made readers distort subsequent contextual information. However, this category was a little different in the sense that the present participants gave literal interpretations even when they were aware of the semantic inconsistency. In the case of

Elaboration of context, the participants resolved the semantic inconsistency by inferring or adding an underlying scenario or explanation that allowed the conflicting morphological and contextual meanings to co-exist. Such elaborative processing also would reflect their mental effort to retain the original text information without changing its literal interpretation. These kinds of changes in literal interpretations of context were found in L2 reading research that examined how readers resolved inconsistent information during reading (Ushiro et al., 2018) and in their memory (Ushiro, Mori, et al., 2016).

Furthermore, the individual patterns of literal interpretation of opaque targets were analyzed in order to explore whether such inferences were the result of problems in learners' inferential skills. The results suggest that none of the 12 participants made literal interpretations frequently (0 to 4 times out of 16 opaque target words). In addition, such inferences were reduced as the number of encounters increased; most of these inferences were made at the beginning of the task. In this regard, it was highly possible that some kind of *learning effect* occurred regarding the resolution of opaque words during the present experiment. On this point, the learning may be the result of increased experience in inferring opaque words rather than an improvement in their inferential skills. In the beginning, they might not know how to resolve such inconsistencies but later learned to avoid such inferences and make appropriate interpretations, rather than making literal interpretations through such strategies as a Persistent reliance on morphemes.

Moreover, it was possible that a lack of readers' perception or metalinguistic knowledge that the morphemes of a word do not always contribute to the word meaning was the primary source of their initial literal interpretations. GEM presupposes that readers' search for actual (context-based) meaning is initiated with their accurate understanding of the surrounding context and perception of semantic inconsistency between the two sources (Levorato & Cacciari, 1995, 1999). On this point, it was unlikely that the present participants were unable to

understand context meanings, considering the material characteristics of this experiment. Therefore, the persistence observed in this study might have come from another source. Such metalinguistic knowledge of the discrepancy between the parts and the whole of the word develops with increased exposure (Cain et al., 2009). The continued encounters with opaque words made them aware of this and they learned to search for context-based meanings, instead of persisting or relying on the meanings of the individual morphemes. However, this does not mean that they had not acquired such knowledge before the experiment because such discrepancies were to some degree pervasive both in L1 and L2, and it is unlikely for them to have acquired such a level of metacognition during the course of the experiment.

Based on these considerations, it can be argued that EFL learners at the intermediate level should be able to make contextually appropriate interpretations of opaque words once they have perceived the semantic inconsistency between morphemes and context, and at the least they should be asked to prioritize context-based meanings for interpretations.

3.3.5 Conclusion of Experiment 3

Experiment 3 was conducted to explore (a) inference performance on opaque words in a lexical inference task, (b) its relationship with lexical inference strategy use, and (b) causes of incorrect word-based interpretations of opaque unknown words.

The main findings of this experiment are as follows. First, EFL learners generally made appropriate interpretations of unknown words, even for opaque words, when they were asked to infer the meanings (RQ3-1). They used various inferential strategies differently, from identifying to metacognitive processes, based on the semantic transparency of unknown words; the use of these strategies contributed to deriving contextually appropriate meanings. For opaque words, EFL readers metacognitively responded to the semantic conflicts of the two

sources, which triggered their effective use of parts of morphemes and context-based inferences for appropriate interpretations (RQ3-2).

Moreover, qualitative analysis enabled us to characterize five processes that led to morpheme-based interpretations of opaque words. While some of these inferences were caused by participants' lack of monitoring of contextual information (Negligence of context, Lack of evaluation), others were made even after perceiving the semantic inconsistency between the two sources: the participants persisted on literal interpretations (Persistent reliance on morphemes) or changed the contextual interpretation (i.e., Modification of context, Elaboration of context). This variety of causes demonstrates the importance of observing the inferential process, which cannot be examined only through inference outcomes (RQ3-3).

In addition, the analysis of individual patterns would indicate that these wrong inferences were not necessarily due to their lack of inferential skills. Rather, it was a problem in their metalinguistic perception of how to deal with unknown words when interpretations based on their constituent morphemes were inconsistent with the surrounding context.

3.4 Conclusion of Study 1

The three experiments of Study 1 investigated intermediate-level Japanese EFL readers' understanding of unknown words with semantically transparent and opaque morphological structures. Specifically, we were interested in determining (a) what information they use for interpretation, (b) the effects of inappropriate morpheme-based interpretations on text comprehension, and (c) the cognitive processes involved in appropriate interpretations.

The on-line reading time data of Experiments 1 and 2 indicated that they showed stable sensitivity to the semantic relationship between morphemes and contexts, even in neutral contexts. However, generating context-based meanings for opaque targets was demanding during reading: the participants often made inappropriate morpheme-based interpretations.

Moreover, these interpretations had some negative impacts on their text comprehension. These misinterpretations are less likely to be revised as a reader proceeds further in a text and can sometimes distort the interpretations of other parts of a text. However, in Experiment 3, such incorrect inferences decreased by asking them to infer the target words explicitly. The lexical inference task solicited their conscious efforts to achieve contextually appropriate interpretations by changing their inferential strategies according to the semantic transparency of the morphemes.

In conclusion, the successful interpretation of opaque words during reading requires a variety of metacognitive processes and a flexible use of linguistic clues, and it is challenging even for intermediate-level learners during reading. However, when it comes to the lexical inferencing skills of these learners, this problem could be trivial in the sense that they could avoid inappropriate inferences if they were given instruction on dealing with unknown words; they would gain the necessary linguistic and metacognitive skills to resolve the semantic inconsistency of morphemes and context.

Chapter 4

Study 2: Use of Morphological and Contextual Information in Lexical Inferencing by Beginner-Level Learners

The three experiments in Study 1 investigated intermediate-level Japanese EFL learners' processing and interpretation of transparent and opaque unknown words. The overall results showed that they were sensitive to the semantic relations between morphemes and context during reading, which is a prerequisite to the successful inferencing of opaque words. However, learners often interpreted opaque words literally although the meanings were semantically inconsistent with the surrounding context. Additionally, literal interpretations sometimes caused individuals to distort the contextual information so that literal interpretations made sense.

However, such inappropriate inferences were reduced when EFL learners were explicitly asked to infer the words, which shows the importance of motivation in generating contextually appropriate meanings. Additionally, although literal interpretations were rarely found, most of these inferences were possibly due to participant perceptions concerning how to resolve semantic conflicts rather than a lack of linguistic skills among participants. Accordingly, the requirements for appropriate interpretation were (a) having participants focus on unknown words and meanings and, if any existed, (b) directing them to prioritize contextually appropriate meanings when words were semantically opaque by using parts of morphemes or context-based meanings and not changing context interpretations.

However, it is premature to conclude that EFL learners in general could process and interpret information in a similar way given the wide variety of individual differences. Among these differences, learner proficiency is the most influential factor (e.g., Bengeleil & Paribakht, 2004; Hamada, 2014; Huckin & Coady, 1999; Mori, 2002). Past studies have shown that less proficient learners often make incorrect word-based inferences (Hamada, 2014; Mori & Nagy,

1999; Ushiro et al., 2010). Thus, we predicted that less proficient EFL learners, particularly at the beginner level, would have more difficulty using multiple sources of information to interpret unknown lexical items.

To date, successful lexical inferencing is considered a demanding task for beginner-level learners because of their limited processing skills in a target language (Huckin & Coady, 1999; Nation, 2013). Thus, some experts insist that asking beginner-level learners to infer by themselves is not a good strategy (Folse, 2004). However, given the multiple processes involved, it is still important to identify what individuals can and cannot do by themselves and what type of support they need for appropriate interpretations. These insights will have implications when determining the focus of instructions when dealing with unknown words in a text.

Accordingly, the primary purposes of Study 2 are (a) to identify the processes by which beginner-level EFL learners have difficulty appropriately interpreting unknown words and (b) explore effective intervention based on the difficulties. Thus, Study 2 focuses on performance in the case of lexical inferencing, not normal reading. This is because the results of Study 1 showed that appropriate interpretation of opaque words while reading was difficult even for intermediate-level learners. Therefore, Experiment 4 attempted to reveal the difficulty involved in appropriate interpretation, and Experiment 5 examines the effects of inference training based on the findings of Experiment 4.

4.1 Experiment 4: Understanding Context-Morpheme Semantic Relations Among Beginner-Level Japanese EFL Learners

4.1.1 Purposes, overview, and research questions of Experiment 4

The purpose of Experiment 4 was to reveal which of the processes involved in the appropriate interpretation of opaque unknown words beginner-level have difficulty in achieving successful inferences. There are two primary processes required for the appropriate

interpretation of opaque words: (a) the initial perception of the semantic relationship between the morphemes and context, and (b) the subsequent generation of context-based meanings.

To this end, Experiment 4 focused only on the former process, perception of the semantic relationship between the two sources, for the following reasons. First, past studies have shown that successful inferencing using both morphemes and context is difficult for learners even when multiple options are available (Hamada, 2014; Mori & Nagy, 1999). Therefore, we expect that generating meanings by themselves would be more difficult when individuals are required to produce their own inferences in an open-ended format as in this study. Second, readers with limited skills tend to read text in a piece-by-piece manner, which could prevent the consideration of both context and word-based meanings (Oakhill et al., 2016), which results in failure to notice the semantic conflicts between the two sources. Therefore, in Experiment 4, we decided to focus only on participants' performance in the first process, in detail, rather than examining both processes. The findings of this study reveal which processes we should focus on when supporting individuals in Experiment 4.

In the experiment, we assessed the participants' understanding of the semantic relation via the semantic consistency judgment task (cf. Koda, 2000). In this task, participants read a single-sentence context with its final word (target word) deleted and were then asked to explicitly judge whether the target words fit the context semantically. The correct response rates and reaction time data for the judgment were used to assess the accuracy and efficiency of participants' understanding of the semantic relationship between the two sources. To closely examine the nature of this processing, we used both on-line and off-line versions of the task.

Additionally, we were also interested in whether beginner-level learners could use contextual information to narrow down or activate upcoming word meanings for Experiment 5. This is because appropriately inferring opaque words in this study requires readers to infer context-based meanings as well as understand the context. To this end, we analyzed the reaction

times for the on-line task. If the participants could use contextual information for inference, the reaction times should become shorter when the informative contexts precede the target word compared to when neutral contexts precede the target word due to the richer semantic information of unknown words derived from informative contexts. As a result, RQ2 addressed in Experiment 4 are as follows.

RQ 4-1: To what degree do beginner-level Japanese EFL learners understand the sematic relationship between morphological and contextual information?

RQ 4-2: Are beginner-level Japanese EFL learners able to narrow down the meanings of unknown words according to context informativeness?

Regarding RQ 4-1, past studies have reported that less skilled readers often fail to notice the semantic inconsistency between morphemes and context because their limited processing skills prevent them from considering or monitoring the two sources (Hamada, 2014; Oakhill et al., 2016). However, the semantic consistency judgment task employed in this experiment explicitly required readers to judge the semantic in/consistency between the two sources. Thus, it was possible that their attention was more directed to both types of information and correctly judged the relations.

As for RQ 4-2, the past finding that less proficient readers can narrow down meaning only to a general category from both strong and week constraining contexts (Hamada, 2013) indicates the difficulty inherent in making use of contextual information. Moreover, given that the present participants are less skilled than the less-proficient group used for Hamada's study, the present participants may not activate more concrete semantics from informative contexts relative to neutral contexts.

4.1.2. Method

4.1.2.1 Participants

The participants of Experiment 4 were 21 Japanese undergraduate students from a Japanese university. None of the students had participated in any other experiments. The major for all students was international business management, and all of them had learned English for at least six years in Japanese educational settings. Their estimated English proficiency was approximately at the beginner level based on self-reported scores on English proficiency tests of EIKEN test (3rd grade: n = 9; pre-2nd grade: n = 4).

As in Study 1, participants' reading proficiency was assessed with a reading proficiency test as a component of this experiment. However, because of the lower proficiency level of the present participants compared to Study 1, the test was adapted from lower grades of the EIKEN test. The test was composed of two passages from the third grade (k = 8), two from the pre-second grade (k = 8), and one from the second grade (k = 4) reading subsection of the EIKEN test (Obunsha, 2000, 2001a, 2001b). The time allocation was identical (20 minutes) to allow a comparison with the results from other experiments. The average score of the reading proficiency test was 9.71 (SD = 2.12) and ranged from 6 to 14 (Cronbach's α = .81).

4.1.2.2 Materials

Target word and context

The materials used in this study were identical to 32 target compounds and paired single context sentences used in Experiment 1 (see Appendix 3). Again, the target words were pseudo compound words (e.g., *pricewar*) whose surrounding context, either informative or neutral, were semantically consistent (transparent) or inconsistent (opaque) with the literal interpretations of target words.

Then, four presentation sets, in which every target word appeared only once for each condition, were created to counterbalance the target words' presentation conditions across participants. Thus, all the participants would read each of the 32 target words in one of the four conditions and encounter eight target words per each of the four conditions.

On-line semantic consistency judgment task

We assessed the participants' on-line understanding of the semantic relation between the two sources using the on-line semantic consistency judgment task (Koda, 2000). The task was conducted on a PC. In this task, the context sentence with the final word deleted was presented. Then, the target word would appear on the next screen, and participants were asked to judge if the target word was semantically consistent ("Yes") or inconsistent ("No") with the preceding context sentence as accurately and quickly as possible. The correct response rates and reaction times were used for analysis. We randomized the presentation order of the target words

Off-line semantic consistency judgment task

To examine the participants' context-target integration under no time limit, we prepared an off-line semantic consistency judgment task (see Figure 4.1). The task was paper-based and, unlike the on-line phase, participants were allowed to take as much time as possible.

1. Peter studied the reason for the **businessbag**.					
太字の単語から考えられる意味は周りの文と意味的に一貫していますか？【 **Yes** ・ **No** 】(どちらかに○)					
1 全く自信がない	2 自信がない	3 あまり自信がない	4 少し自信がある	5 自信がある	6 とても自信がある

Figure 4.1. An example of the off-line semantic consistency judgment task in Experiment 4.

They were asked to judge the target words' fitness to the context by circling either yes or no on the sheet. Additionally, participants were asked to assess their confidence in their judgments using a six-point Likert scale (1: *not confident at all*; 2: *unconfident*; 3: *a little unconfident*; 4: *a little confident*; 5: *confident*; 6: *very confident*).

The 32 target-context pairs for the off-line task were identical to the on-line task for each participant. This allowed us to directly compare the effectiveness of morpheme-context integration by task conditions (i.e., on-line vs. off-line). Although this meant that the participants had already been exposed to the same materials by the time they worked on the off-line task, there appeared to be no clear advantage for the performance of the off-line task because prior exposure to materials would not render the off-line judgments easier where a faster decision was not required.

4.1.2.3 Procedure

The experiment was conducted individually in a silent room. First, the author explained the general purpose and the entire experiment procedure. Then, the participants were randomly assigned to one of the four presentation sets.

Next, participants worked on the two semantic consistency judgment tasks. The on-line task was administered using SuperLab 5.0 and Response Pad RB-740. Each trial began with a "Ready?" screen. When participants were ready, they were instructed to push the yes button. The next screen showed the context sentence with the final word (target word) deleted. When they finished reading the screen, participants were asked to push the button again. Pressing the button caused the target word to appear, and the participants were then asked to judge whether or not literal meanings of morphemes of target words fit to the preceding context semantically as accurately and quickly as possible by pressing either yes or no. Although there was pressure to make a quick decision, a time limit was not set (Koda, 2000). Before the experimental session,

the participants took three practice trials to familiarize themselves with the procedure. After the trials, the participants performed the judgments for the 32 pairs.

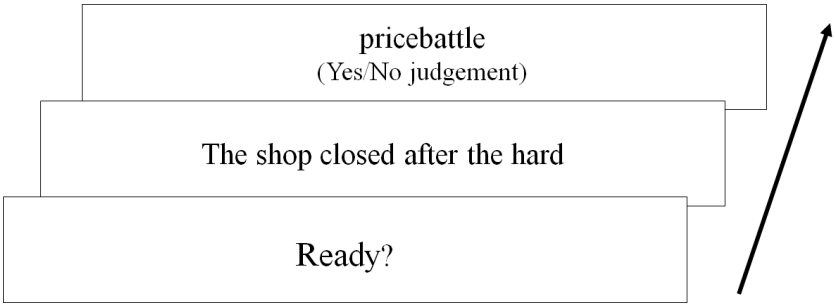


Figure 4.2. An example of the on-line semantic consistency judgment task in Experiment 4.

After a short break, the participants proceeded to the off-line task. In this task, the same 32 morpheme-context pairs with a randomized order of presentation were presented in a booklet. The participants were asked to judge the semantic fitness of the target words to the context and their level of confidence with those judgments on a six-point Likert scale. At that time, unlike the on-line phase, complete sentences including target words were available. The participants were allowed to take as much time as they needed to complete the tasks. It took approximately 20 to 35 minutes to complete the two tasks. After the task, they were notified of the use of pseudowords for target words. Lastly, the participants took the English reading proficiency test, which took the participants approximately 20 minutes to complete.

4.1.2.4 Analysis

The data from all participants were submitted for analysis because they showed relatively high performance based on the correct response rates (On-line: M = 77.57%, 95% CI [74.45, 80.68], SD = 6.83, Min = 66.52%, Max = 90.63%; Off-line: M = 80.37%, 95% CI [77.64, 83.11], SD = 6.00, Min = 72.42%, Max = 90.63%).

Correct response rates

The correct response rates for both the on- and off-line semantic consistency judgment task for transparent and opaque target words were calculated, respectively. For transparent words, the correct response was yes, and no was the correct response for opaque words. To explore their performance differences, we conducted a 2 (transparency: transparent, opaque) 2 × 2 (context: informative, neutral) × 2 (task: on-line task, off-line task) three-way ANOVA on the correct response rates; all three variables were within-participant designs.

Reaction times

We used the reaction time data for the on-line task to examine the efficiency of the judgments. Before the main analysis, we treated the reaction time outliers. Reaction times above three standard deviations for each participant within the same cell were considered outliers (Koda 2000). Additionally, we also discarded trials where participants skipped either context or target words. Reaction times under 200ms were considered to be such data. Overall, 5.6% of the total data were discarded.

All the valid reaction time data, irrespective of judgment accuracy, were submitted for analysis (Koda, 2000). We ran a 2 (transparency: transparent, opaque) and 2 (context: informative, neutral) repeated ANOVA on the participants' reaction times using transparency and context as within-participant variables.

4.1.4 Results

4.1.3.1 Correct response rates

The descriptive statistics of the correct response rates for the on- and off-line semantic consistency judgment tasks are presented in Table 4.1 and graphically in Figure 4.3. Overall, the correct response rates were higher, irrespective of conditions and context informativeness.

Table 4.1

Descriptive Statics of Correct Response Rates (%) for On- and Off-line Consistency Judgments in Experiment 4

	On-line task				Off-line task			
	Informative		Neutral		Informative		Neutral	
Transparency	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Transparent	76.96	16.16	69.39	16.79	86.31	11.79	79.76	12.12
Opaque	83.65	12.61	80.27	12.57	83.93	14.94	82.74	15.86

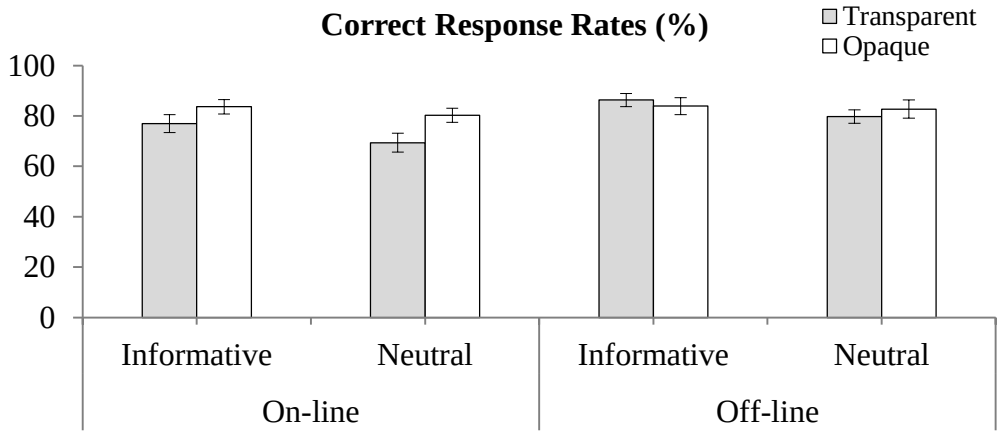


Figure 4.3. Correct response rates for the on- and off-line semantic consistency judgment tasks in Experiment 4.

We conducted a 2 (transparency) × 2 (context) × 2 (task) three-way ANOVA on the correct response rates (see Table 4.2). The results showed that the main effect of task reached significance, $F(1, 20) = 12.34$, $p = .002$, $\eta_p^2 = .382$, and the main effect of context was marginally significant, $F(1, 20) = 3.28$, $p = .085$, $\eta_p^2 = .141$. Additionally, Task × Transparency interaction was significant, $F(1, 20) = 6.70$, $p = .018$, $\eta_p^2 = .251$. On the other hand, the main effect of transparency, $F(1, 20) = 2.34$, $p = .142$, $\eta_p^2 = .105$, and other interactions were all insignificant (*p*s > .10).

Table 4.2

Summary Table for Three-Way ANOVA of the Effects of Task, Context, and Transparency on Correct Response Rates in the On- and Off-line Semantic Consistency Judgment Task in Experiment 4

Source	*SS*	*df*	*MS*	*F*	*p*	η_p^2
Task (Ta)	1326.21	1	1326.21	12.341	.002	.382
Error (Ta)	2149.29	20	107.46			
Context (C)	915.79	1	915.79	3.277	.085	.141
Error (C)	5588.88	20	279.44			
Transparency (Tr)	866.50	1	866.50	2.337	.142	.105
Error (Tr)	7415.60	20	370.78			
Ta × C	26.91	1	26.91	0.168	.687	.008
Error (Ta × C)	3211.38	20	160.57			
Ta × Tr	756.67	1	756.67	6.704	.018	.251
Error (Ta × Tr)	2257.43	20	112.87			
C× Tr	239.53	1	239.53	0.815	.377	.039
Error (C × Tr)	5878.41	20	293.92			
Ta × C × Tr	3.54	1	3.54	0.039	.846	.002
Error (Ta × C × Tr)	1821.14	20	91.06			

To interpret the Transparency × Task interaction, we performed post hoc tests. The results revealed that correct response rates were higher for transparent words from the off-line task than transparent words from the on-line task ($p < .001$) while correct response rates for opaque words did not differ by task ($p = .567$). Additionally, correct response rates for opaque words were higher than correct response rates for transparent words from the on-line task ($p = .010$) while no difference was observed for the off-line task ($p = .928$).

The overall results are summarized as follows. First, performance regarding transparent words from the off-line task was superior compared to performance regarding transparent words from the on-line task. Second, correct answer rates were higher when the informative context preceded the target words. Lastly, for the on-line task, opaque words were judged more correctly than transparent words.

4.1.3.2 Reaction Times

Table 4.3 shows the descriptive statistics of the reaction times for the on-line semantic consistency judgment task for each condition, and Figure 4.4 shows the statistics graphically. Regardless of the semantic transparency of target words, reaction times seem faster for informative contexts.

Table 4.3

Descriptive Statistics of Reaction Times (ms.) for the On-line Consistency Judgment Task in Experiment 4

	Informative context			Neutral context		
Transparency	*M*	95% CI	*SD*	*M*	95% CI	*SD*
Transparent	3154.82	[2706.39, 3603.26]	985.15	3671.92	[3322.93, 4020.90]	766.67
Opaque	3338.26	[2873.17, 3803.35]	1021.73	3614.21	[3153.88, 4074.54]	1011.28

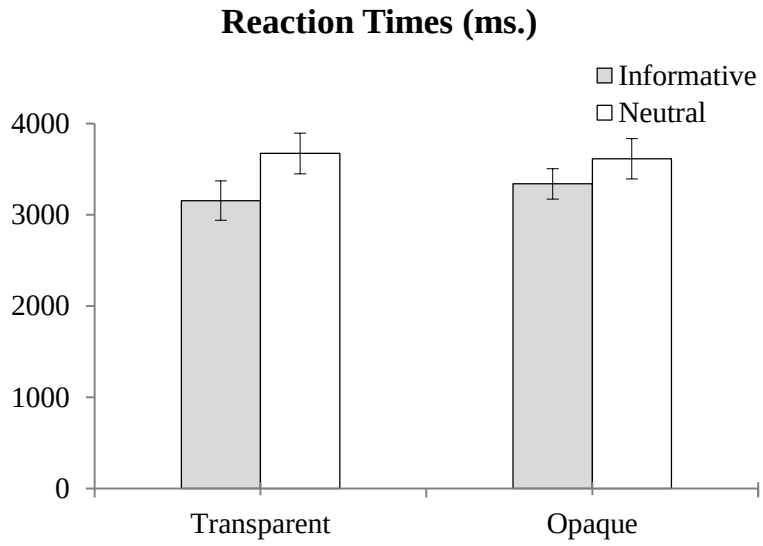


Figure 4.4. Reaction times for the on-line semantic consistency judgment task in Experiment 4.

To explore the factors affecting the decision as to the semantic consistency, we performed 2 (transparency) × 2 (context) two-way ANOVA (see Table 4.4). The results showed that the main effect of context was significant, $F(1, 20) = 5.31$, $p = .032$, $\eta_p^2 = .210$, which suggests that the participants responded faster when targets were preceded by an informative context rather than a neutral context. However, neither the main effect of transparency, $F(1, 20) = 0.22$, $p = .645$, $\eta_p^2 = .011$, nor Transparency × Context interaction, $F(1, 20) = 1.59$, $p = .222$, $\eta_p^2 = .074$, were significant.

Table 4.4

Summary Table for Two-Way ANOVA of the Effects of Transparency and Context on Reaction Times for the On-line Semantic Consistency Judgment Task in Experiment 4

Source	*SS*	*df*	*MS*	*F*	*p*	η_p^2
Transparency (T)	82996.75	1	82996.75	0.219	.645	.011
Error (T)	7563878.62	20	378193.93			
Context (C)	3301793.15	1	3301793.15	5.308	.032	.210
Error (C)	12440182.75	20	622009.14			
T × C	305292.37	1	305292.37	1.590	.222	.074
Error (T × C)	3840173.94	20	192008.70			

4.1.3.3 Confidence Ratings

Finally, Table 4.5 shows the descriptive statistics for the confidence ratings for the off-line task. Overall, the value was approximately 4 [*a little confident*], irrespective of the conditions.

Table 4.5

Descriptive Statistics of Confidence Ratings for the Off-line Consistency Judgment Task in Experiment 4

	Informative context			Neutral context		
Transparency	*M*	95% CI	*SD*	*M*	95% CI	*SD*
Transparent	4.15	[3.75, 4.55]	0.88	3.85	[3.40, 4.31]	1.00
Opaque	3.87	[3.59, 4.15]	0.61	3.71	[3.30, 4.13]	0.91

Note. Range = 1 [*very unconfident*] to 6 [*very confident*].

A 2 (transparency) × 2 (context) two-way ANOVA (see also Table 4.6) showed the main effect of context, $F(1, 20) = 9.68$, $p = .005$, $\eta_p^2 = .326$. Neither the main effect of transparency, $F(1, 20) = 2.79$, $p = .110$, $\eta_p^2 = .122$, nor Transparency × Context interaction, $F(1, 20) = .017$, $p = .899$, $\eta_p^2 = .001$, reached statistical significance.

In summary, the confidence rating was affected by the context informativeness of the preceding context: The participants became increasingly confident in their judgments when the surrounding context was informative compared to when the surrounding context was neutral.

Table 4.6

Summary Table for Two-Way ANOVA of the Effects of Transparency and Context on Confidence Ratings for the Off-line Semantic Consistency Judgment Task in Experiment 4

Source	*SS*	*df*	*MS*	*F*	*p*	η_p^2
Transparency (T)	0.34	1	0.34	2.06	.167	.093
Error (T)	3.30	20	0.17			
Context (C)	2.17	1	2.17	11.425	.003	.364
Error (C)	3.80	20	0.20			
T × C	0.02	1	0.02	0.125	.727	.006
Error (T × C)	3.76	20	0.19			

4.1.4 Discussion

To what degree do beginner-level Japanese EFL learners understand the sematic relationship between morphological and contextual information? (RQ 4-1)

The results of the two types of semantic consistency judgment tasks showed that the participants' responses were typically correct irrespective of context informativeness and task conditions (69.4–86.3%). Although there may be some performance differences according to some factors, the overall results suggest that EFL learners, even at the beginner level, could understand the semantic relationship between morphemes and context when they were asked to judge the relationship.

This is partly inconsistent with previous findings whereby less skilled readers failed to monitor the semantic relation of multiple information, which results in incorrect literal interpretations (Cain et al., 2009; Hamada, 2014; Oakhill et al., 2016). The piece-by-piece reading style typical of less skilled readers make it difficult for them to consider context-based meaning (Oakhill et al., 2016).

There are two possible reasons for this discrepancy. First, Hamada's (2014) discussion is based on the outcomes (selected answer options) of lexical inferencing, not on observations of the inferential process. Thus, failure to make accurate inferences in her study might reflect readers' inability to make appropriate inferences and not on a failure to judge semantic inconsistencies. This is partly consistent with a homonym by Ushiro et al. (2010) in which some participants stuck to the known word meanings even when they could accurately comprehend the text because of their resistance to changing interpretations. Similarly, such resistance was also observed in the think-aloud comments from Experiment 3, in which the participants were aware that context-based meanings were inconsistent with the surrounding context but used literal interpretations (i.e., Persistence reliance on morphemes).

The second reason is related to the degree of participants' attention to the two sources. In

a lexical inferencing task, participants are asked only to infer the unknown words' meanings; thus, their goal was to generate the target word meanings or choose from answer options in the case of a multiple-choice format. Therefore, it was possible that the participants could complete the required products (inferred word meaning) without necessarily considering the two sources. On the other hand, as in this experiment, the yes-no response involved in the judgment directed their attention to the semantic consistency of morphemes and context. Therefore, even beginner-level learners could understand the relationship if their attention is strongly directed toward it, as in this experiment.

However, despite the overall high accurate response rates for the consistency judgments, participants' performance was affected by factors such as task mode and context quality. We discuss these effects and explore the nature of beginner-level learners' perceptions of consistency between the two sources.

The ANOVA results on response accuracy showed that the significant Task × Transparency interaction suggests that correct answer rates for transparent words were higher for the off-line task than the on-line task ($p < .001$), and performance with opaque words was higher for transparent words than opaque words ($p = .010$). Collectively, these results indicate that although performance is lower for the on-line task, this was seen to be the case only for transparent words.

This performance difference for transparency is strange because if the readers could comprehend the context messages appropriately, the in/consistency between morphemes and context should be perceived similarly as in the norming phase of the material development of this study. Therefore, this raises the possibility that participants might have made judgments relying on, to some degree, imperfect or surface representations of context. To explore this possibility, we discuss it in terms of the cognitive process involved in correct responses for transparent words and the online task along with other present data.

For opaque words, the correct response (“No”) required participants to understand whether literal interpretation of the target words did not fit the context. Thus, it may be possible that they could correctly judge once they understood that morpheme-based meanings were semantically irrelevant to the general topic of the surrounding context. Since the literal meaning of opaque words in this study were unrelated or contracted with the context meanings, the semantic inconsistency was somewhat salient. As a result, even if their understanding of the context was not accurate, participants could correctly detect the semantic inconsistencies with partial contextual information. On the other hand, for the transparent condition, a correct response (“Yes”) would require a more accurate representation of the entire context to reject the possibility that the literal meaning did not fit semantically.

Additionally, for the on-line task, participants had to rely on the sentence representations available in their working memory because when the target words were presented, the context disappeared. On the other hand, for the off-line task, both target words and context were available. Thus, participants were able to check the relations more carefully, and the accurate response rates did not differ by semantic transparency. Accordingly, we could argue that the above two factors additively impaired the correct response rates for the transparent words from the on-line task.

This is also consistent with the marginal significance of the main effect of context: the correct response rates were slightly higher for the informative contexts than the neutral contexts. Although the semantic transparency of a word resides on a continuum from fully opaque to fully transparent (Libben & Jarema, 2006; Schäfer, 2018), whether a certain word belongs to the transparent or opaque side can be dichotomous. Therefore, consistency judgment performance should not be affected by context informativeness if readers accurately understand the context. Accordingly, the responses with higher accuracy for informative contexts conformed to the possibility that their judgment was made based on surface representations of

context rather than cumulative meanings of context.

Moreover, this possibility is also supported by the low level of confidence participants showed in their judgments for the off-line task. The overall confidence rating was approximately 4 (i.e., "*a little confident*") irrespective of conditions (3.71–4.15). Additionally, participants' confidence was higher for informative contexts than neutral contexts, which converges on the correct response rates.

In summary, the correct response rates results suggest that Japanese EFL learners, even at the beginner level, understand the semantic relationship between morphemes and context. However, we also noted that they might have relied on the judgment based on surface or imperfect contextual representations.

Are beginner-level Japanese EFL learners able to narrow down the meanings of unknown words according to context informativeness? (RQ 4-2)

To examine the efficiency of consistency judgment and participants' ability to exploit contextual information for lexical inference, we analyzed the reaction times for the on-line semantic consistency judgment task. The ANOVA results suggested that only the main effect of context was statistically significant: the reaction times were shorter for informative contexts than neutral contexts. Lack of Transparency × Context interaction showed that the trend was observed for both transparent and opaque words. More concrete semantics derived from the informative context made their judgments faster than those in neutral contexts where abstract meanings can be inferred. Thus, even beginner-level learners can activate or predict the semantic representation of upcoming words from context according to the context informativeness.

However, this result contradicts a previous finding that reported the effects of reading proficiency on the degree of inference specificity (Hamada, 2013). In Hamada's study, lower-

proficiency EFL readers could activate only a general category of words from weakly constraining context and strongly constraining contexts when reading while high-proficiency learners could activate more specific meanings from strongly constraining contexts.

The possible reason is attributed to the different task nature of normal reading and the semantic consistency judgment of this study. In the on-line task, the participants read the context so that they could judge whether the meaning of the context was semantically consistent with the subsequently presented target words. Therefore, it is highly possible that they attempted to read the context by predicting the semantic information of the upcoming words during the context reading. However, what type of semantic representation of upcoming words the participants constructed from informative contexts was unclear; therefore, based on the inference outcomes observed in Experiment 5, further examination is needed.

4.1.5 Conclusion of Experiment 4

The purpose of Experiment 4 was to examine participants' understanding of the semantic relationship between the two sources and their ability to use contextual information for inferences. In the experiment, participants judged the semantic in/consistency between morphemes and context via both on- and off-line semantic consistency judgment tasks.

The correct response rates showed that beginner-level learners could understand the semantic relationship between the two sources: all participants correctly judged the semantic in/consistency of morphemes and context in most cases (On-line: 77.6%; Off-line: 80.4%). However, lower performance for transparent words in the on-line task indicated that their judgment was based on the superficial representation of context.

The on-line reactions time showed that participants could predict the semantic information of upcoming words from contextual information. This was suggested by the shorter reaction times when the preceding context had informative information for words compared to

neutral contexts.

These findings have some implications for Experiment 5, in which we examine the effects of inference training. First, the intervention should primarily focus on how to generate context-based meanings since participants understood the semantic relation by themselves when their attention was directed to the relation. Second, training materials are expected to use only informative contexts. The participants of this experiment could predict more concrete semantics from informative contexts; thus, directive contexts can be used as a suitable first step for beginner-level learners given the difficulty of generating contextually appropriate meanings for intermediate-level learners in Study 1.

4.2 Experiment 5: Effects of Lexical Inference Training on the Use of Morphological and Contextual Information

4.2.1 Purposes, overview, and research questions of Experiment 5

The results of Experiment 4 showed that even beginner-level EFL readers could understand the semantic relationship between morphemes and their context when they were explicitly asked to judge it, which is a necessary first step in correctly inferring the meaning of unknown opaque words. However, given the difficulty of inferring unknown words, especially opaque words for less-skilled learners (Hamada, 2014), it is highly likely that more help will be required in the subsequent process, that is, generating contextually appropriate interpretations. Therefore, Experiment 5 examined the effects of inference training aimed at guiding the learners to generate contextually appropriate meanings. The training employed an inferencing procedure that was developed by the author, based on the findings of the previous experiments of this study. The findings of this experiment will further our understanding of the kind of support that is necessary for beginner-level learners to engage in successful lexical inferences.

Concerning an effective inferencing procedure, Clarke and Nation (1980) recommended one consisting of five steps. The detailed procedure is described in Table 4.7. In this approach, after determining the part of speech of the unknown word (Step 1), learners are asked to explore its immediate context (Step 2), guess its meaning (Step 3), check the appropriateness of the guess (Step 4), and evaluate the inferred word meaning while considering morphological information (Step 5).

Table 4.7

Clarke and Nation's (1980) Five-Step Inferencing Procedure

Step 1	Decide on part of speech of the unknown word.
Step 2	Look at the immediate context of the word, that is, the relationship with adjoining sentences or clauses.
Step 3	Guess.
Step 4	Check the guess.
Step 5	Is the guess the same part of speech as the unknown word? Substitute the guess for the unknown word. Does it fit comfortably into the context? Break the unknown word into parts. Does the meaning of the parts support the guess? Look up the word in the dictionary.

The basic idea of this approach is that making morphological analysis the final step of inferencing could help learners avoid incorrect meaning generation without considering contextual meaning. Although the effectiveness of this method has not yet been verified empirically, it appears to reduce incorrect word-based inferences.

However, this approach may be a little impractical and ineffective for the following reasons. First, as Fraser (1999) pointed out, readers' access to word constituents is a somewhat automatic process that they do not have control of; thus, it would be unrealistic for them to consider other clues without processing them. Therefore, the present procedure focused on how to use contextual information after readers' activation of morpheme-based interpretations.

In addition, as found with the differential inferencing strategy use according to the semantic transparency of words in Experiment 3, the minimum cognitive processes involved in appropriate interpretation differ depending on the semantic transparency of the morphemes (see Figure 4.5). For transparent words, due to the semantic convergence of the two sources, it is unnecessary to generate fully contextually-based meanings. What learners must do is perceive the consistency and adopt morphemes-based meanings because they usually convey more concrete meanings than the context (Nagy & Anderson, 1984). For opaque words, however, the semantic conflicts that occur between morphemes and context require more complex processes for appropriate interpretations. In addition to perceiving semantic inconsistencies, readers need to generate contextually appropriate interpretations and suppress unnecessary morpheme-based meanings. Because of these differences, always following the five steps recommended by Nation (2013) might be time consuming and impractical. Generally, readers make minimal efforts to infer unknown words (Hu & Nassaji, 2012; Mondria & Wit-de Boer, 1991; van den Broek et al., 2018), and such a lengthy inferencing procedure would impede fluent reading.

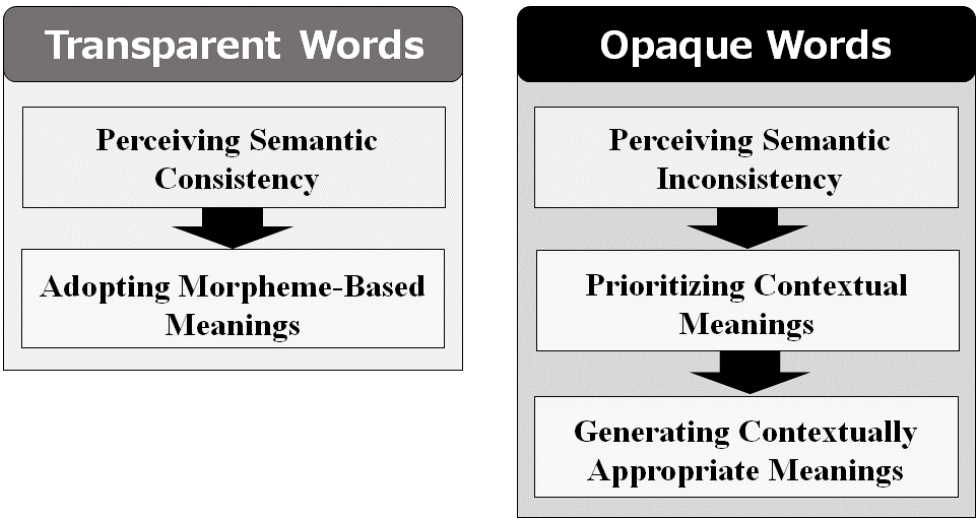


Figure 4.5. Processes involved in appropriate interpretations of transparent and opaque words.

Accordingly, the present study used a procedure that could not only avoid incorrect word-based interpretation but also required a minimal amount of effort according to the semantic transparency of unknown words. To do so, we drew on the findings of the previous experiments of this study, as described below. Figure 4.6 presents the cognitive processes involved in the appropriate interpretation of opaque words and their relationship with the intervention used in the inference training.

First, the results of Experiment 4 suggested that beginner-level EFL learners could perceive the semantic relationship between the two sources of information when their attention was directed to it. Thus, the present study included this process by explicitly asking the learners to judge the semantic relationship between the morphemes and their context.

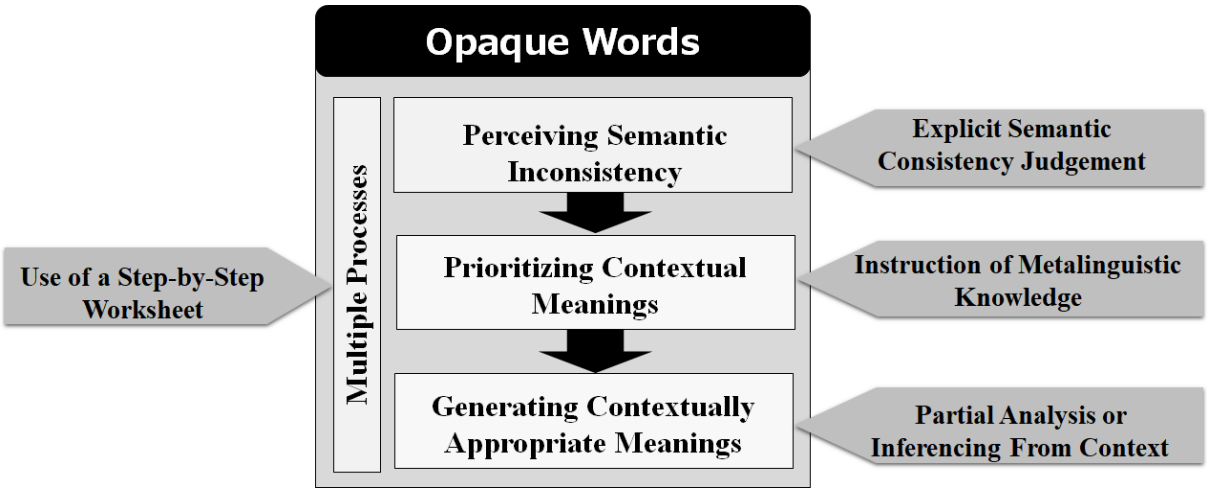


Figure 4.6. Relationship between cognitive processes of generating appropriate interpretations of opaque words and ways of intervention used in Experiment 5.

Second, the results of Experiment 3 indicated that even if the learners could perceive the semantic inconsistency of opaque words, it did not always initiate their search for contextually appropriate meanings. To do this, readers must have the metalinguistic awareness to understand that morpheme-based meanings are not always reliable and thus prioritize context-based

meanings. In this regard, participants will be asked to explore context-based meanings by receiving instructions concerning metalinguistic knowledge about the semantic conflicts between the two sources, and having them prioritize contextual meanings in case of conflict.

In addition, Experiment 3 also identified some strategies that were related to appropriate interpretations of opaque words. Intermediate-level readers used part of morphological information (i.e., Partial analysis) or generated full, context-based meanings (i.e., Inferring from context) for appropriate interpretations. Therefore, the present study supported the meaning-generation process by encouraging learners to use these inferential strategies.

However, given the multiple processes involved in successful interpretation of opaque words, just explaining the procedure to the learners would be too cognitively demanding for beginners to follow, in addition to carrying out each of the above processes. Therefore, to reduce the cognitive burden, the present study used a worksheet that designated when and how they should use linguistic cues to make inferences in a step-by-step manner. This will also serve to identify which steps of the process they had difficulty with.

This experiment employed a pre-test-post-test design. The beginner-level learners completed a onetime lexical inference training, where they inferred unknown words using a worksheet that described how to do so according to the semantic transparency of the morphemes. The participants also took a pre-test and post-lexical inferencing test one week after the treatment, in which the worksheet was not provided to assist them with the process. The participants' performance on the pre-test and the inference training will be compared to see whether they could make more appropriate inferences by using the inference worksheet. In addition, we examined the effects of the training on the participants' normal lexical inferencing, where the worksheet was unavailable. This was done by comparing the results of the pre- and post-tests. The RQs addressed in Experiment 5 are as follows:

RQ5-1: Does lexical inference training using an inference worksheet support the lexical inference performance of beginner-level Japanese EFL learners?

RQ5-2: Does lexical inference training improve the lexical inference skills of beginner-level Japanese EFL learners?

4.2.2 Method

4.2.2.1 Participants

The original participants of this experiment were 149 Japanese first-year university students from a Japanese university from three intact English classes. However, participants who were absent from any of the experimental sessions (n = 17), and who did not follow the procedure of the lexical inference training (n = 13) were excluded; as a result, data from 119 participants were included in the analysis. Their majors were education and international business management. All participants had studied English for at least six years in Japanese formal education. Their estimated English proficiency was approximately beginner level based on their self-reported scores from various English proficiency tests, such as the TOEIC Listening and Reading Test (n = 1, M = 250), and the EIKEN test (4th grade: n = 5; 3rd grade: n = 31; pre-2nd grade: n = 8).

The participants' reading proficiency was assessed using the reading proficiency test (k = 20) used in Experiment 4 (Cronbach's α = .83). Since there was a large variation in their reading test scores (M = 7.10, SD = 2.96, Min = 1, Max = 13), the participants were further divided into three proficiency levels: Upper, Middle, and Lower groups (See Table 4.8). A one-way ANOVA was run to confirm the differences among groups. The results showed a main effect for proficiency, $F(2, 116) = 353.81$, $p < .001$, $\eta_p^2 = .869$. Multiple comparisons with the Bonferroni adjustment revealed significant differences between all groups (all ps < .001).

However, it should be noted that, although there were significant proficiency differences among the three groups, all the participants were considered to have beginner-level English proficiency.

Table 4.8

Descriptive Statistics of the Reading Proficiency Test in Experiment 5

Proficiency	*n*	*M*	95% CI	*SD*	*Min*	*Max*
Upper	37	10.73	[10.22, 11.24]	1.54	9	13
Middle	41	6.93	[6.70, 7.15]	0.72	6	8
Lower	41	4.00	[3.68, 4.32]	1.02	1	5
Total	119	7.10	[6.56, 7.64]	2.96	1	13

Note. Maximum possible score was 20.

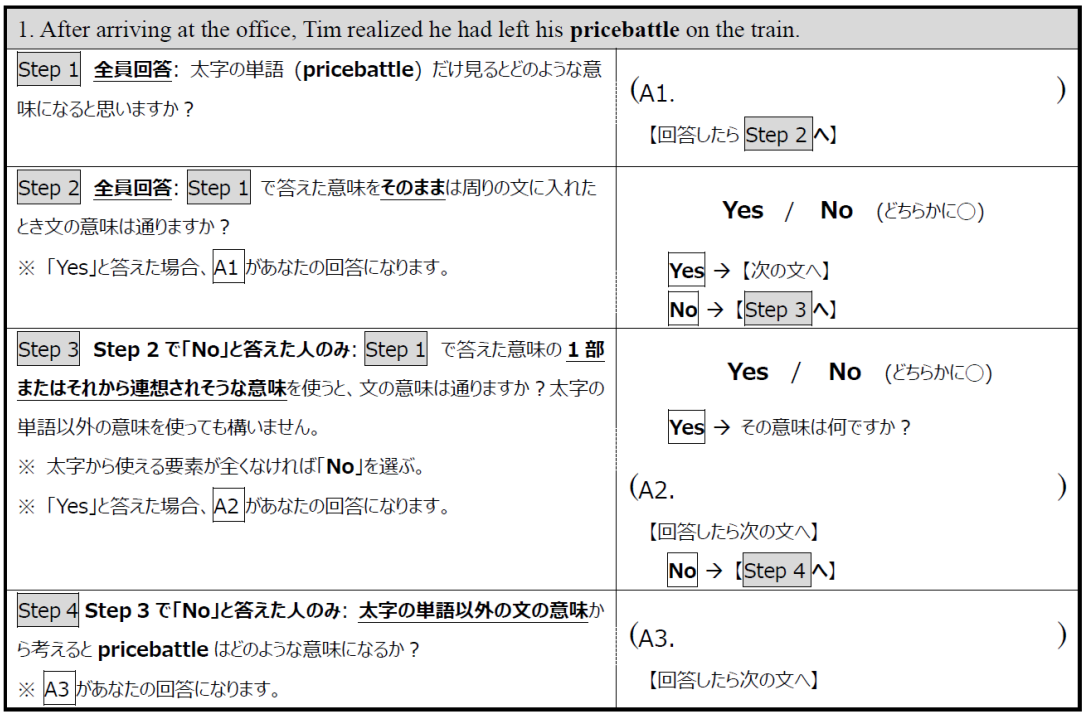
4.2.2.2 Materials

Target Words and Context Sentences

The target words and context sentences were those used in the previous experiments of the present study (Experiments 1, 3, and 4). The 32 target words (e.g., *businessbag*) were either semantically consistent (transparent; e.g., *After arriving at the office, he realized he had left his* ***businessbag*** *on the train*) or inconsistent (opaque; e.g., *After arriving at the office, he realized he had left his* ***pricebattle*** *on the train*) with the surrounding context sentences. Unlike in the other experiments, only informative contexts were used in Experiment 5. Half of the target words (8 transparent and 8 opaque words) were used for the pre- and post-lexical inference tests, and the other half were used for the inference training. Two presentation sets were prepared to counterbalance the effects of the materials' characteristics.

Lexical Inference Worksheet

A lexical inference worksheet was developed by the author for use in the inference training (see Figure 4.7). It consisted of four steps, and the participants were meant to have finalized their inference once they had reached the end of the procedure. Steps 1 and 2 were mandatory for all participants, and Steps 3 and 4 were optional based on their answers in Step 2 and 3. Below are the detailed descriptions of each step.



1. After arriving at the office, Tim realized he had left his **pricebattle** on the train.	
Step 1 **全員回答**: 太字の単語（**pricebattle**）だけ見るとどのような意味になると思いますか？	(A1.) 【回答したら Step 2 へ】
Step 2 **全員回答**: Step 1 で答えた意味を**そのまま**は周りの文に入れたとき文の意味は通りますか？ ※「Yes」と答えた場合、A1 があなたの回答になります。	**Yes** / **No** (どちらかに○) **Yes** → 【次の文へ】 **No** → 【Step 3 へ】
Step 3 **Step 2 で「No」と答えた人のみ**: Step 1 で答えた意味の **1 部またはそれから連想されそうな意味**を使うと、文の意味は通りますか？太字の単語以外の意味を使っても構いません。 ※ 太字から使える要素が全くなければ「**No**」を選ぶ。 ※「Yes」と答えた場合、A2 があなたの回答になります。	**Yes** / **No** (どちらかに○) **Yes** → その意味は何ですか？ (A2.) 【回答したら次の文へ】 **No** → 【Step 4 へ】
Step 4 **Step 3 で「No」と答えた人のみ**: **太字の単語以外の文の意味**から考えると **pricebattle** はどのような意味になるか？ ※ A3 があなたの回答になります。	(A3.) 【回答したら次の文へ】

Figure 4.7. An example of the lexical inferencing procedure using the worksheet provided in the inference training in Experiment 5.

Step 1: Literal interpretations of morphemes (mandatory). In Step 1, all participants were required to come up with literal meanings of the target words based on their two morphemes. This was to reflect situations in which readers have access to the morphological information (Fraser, 1999; Pollatsek et al., 2008, 2011). They were then asked to write the literal meaning down in Japanese in the corresponding blank (A1).

Step 2: Semantic consistency judgment (mandatory). In Step 2, they were asked to judge whether the meanings of the context sentence and the target word were semantically consistent or inconsistent with the literal meanings of the morphemes they produced in Step 1 (A1), as in Experiment 4. The purpose of this step was to evaluate the consistency of the two sources of information, which was an essential step for the subsequent selective use of morphemes according to the semantic transparency of the unknown words. In this step, the judgments were to be made based on whether the sum of the literal meanings of both morphemes of the target word fit the context semantically, not just one of them. The participants indicated their answer by circling either "Yes" or "No" on the worksheet.

If they chose "Yes," the answer produced in A1 (their literal interpretation of the two morphemes) was used as their inferred meaning of the target word, and they were then asked to proceed to the next target word. This was because if the morpheme-based interpretations (A1) fit to the context semantically, the meanings would be not only contextually appropriate but also have more concrete semantics than the contextual meanings. On the other hand, if they answered "No," they moved on to Step 3 because their literal interpretation did not make sense in the context sentence, and they thus had to infer another meaning to obtain an appropriate interpretation.

Step 3: Considering partial use of morphemes (only those who answered "No" in Step 2). The purpose of Step 3 was to consider both morphemes and the context sentence to make contextually appropriate inferences, using some of the morphemes and the target word's context. The participants were asked to consider whether there were any morpheme-based semantic elements that would be semantically possible in the context. They were encouraged to consider partial or associative meanings of morphemes as well as contextual meanings. This process reflects the inferential strategy, Partial analysis, observed in Experiment 3.

If they answered “Yes” (i.e., there were some morphological elements that could be possible in the context), they were asked to write it down on the answer sheet (A2) and then move on to the next target word. In that case, A2 was used as their inferred meaning. On the other hand, if they answered “No,” they were asked to move on to Step 4 because they thought that none of the morpheme-based meanings could be used for appropriate understanding of the target word.

Step 4: Considering context-based meanings (only those who answered “No” in Step 3). The purpose of this step was for the participants to infer the meanings of the target words solely from their surrounding context because none of the morpheme information contributed to the whole-word meaning based on their judgments in the previous steps. At this time, the participants were asked to infer the word meaning solely based on the contextual meaning without using any morphological information, and then write it down on the answer sheet (A3).

In Steps 3 and 4, if they thought that there might be multiple possible semantic options based on the context, they were instructed to write down the first interpretation that they thought of. Note, however, that given the linguistic nature of the materials used in Experiment 5, (a) target words were either transparent or opaque, and (b) appropriate interpretations were judged based on their fitness with the context. One might wonder whether Step 3 (consideration of the partial meanings of morphemes) is unnecessary. However, this was to apply the present findings to other lexical items with multiple interpretations, such as semi-transparent words, (e.g., *honeymoon*), polysemous words (e.g., *lamb*), and idiomatic expressions (e.g., *a piece of cak*e). For these words or phrases, readers can infer accurate meanings using word-based information, such as the primary meanings of polysemous words for peripheral meanings of polysemous words or literal interpretations of the metaphorical meanings of idiomatic expressions.

Questionnaire

To examine the participants' perceived difficulty and task completion of the inference training, a questionnaire was prepared ($k = 3$). The first question (Q1) asked "How difficult was the step-by-step inferencing procedure you practiced last week?" (perceived difficulty of training). The second question (Q2) asked "How well were you able to follow the step-by-step inferencing procedure last week?" (perceived task completion during training). The third question (Q3) asked "Were you able to complete today's task [the post-test] remembering the inferencing procedure you practiced last week?" (perceived task completion during post-test).

All questions were written in Japanese, and the participants were asked to answer the questions by selecting the most suitable response based on a 5-point scale (Q1: 1 [*very difficult*] to 5 [*very easy*]; Qs 2, 3: 1[*not at all*] to 5 [*very well*]).

4.2.2.3 Procedure

The experiment consisted of Sessions 1 and 2 and took place in the students' regular university English classes. Session 2 was conducted one week after Session 1. Figure 4.8 outlines the procedure of this experiment.

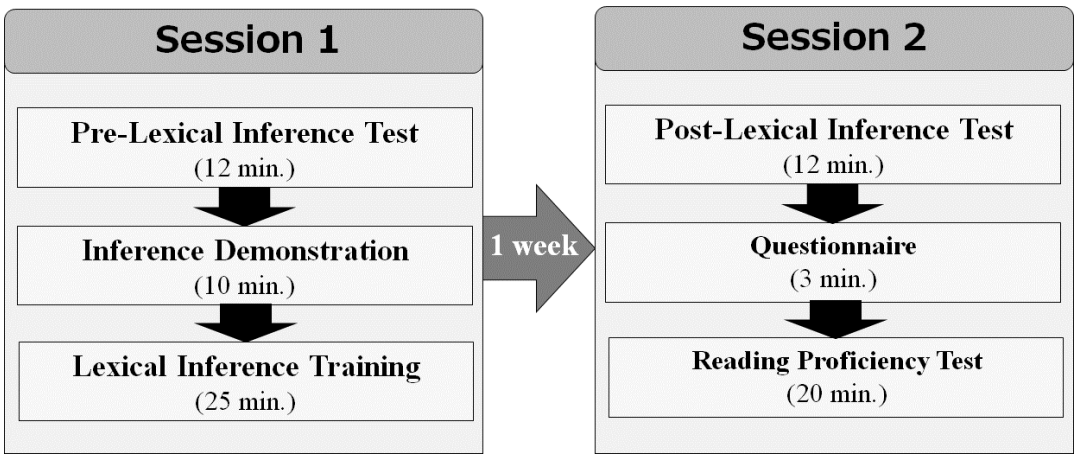


Figure 4.8. The procedure of Experiment 5.

Session 1 consisted of the pre-lexical inference test (pre-test) and the lexical inference training. First, the author explained the general purpose of the experiment, what lexical inferencing is, and its importance in reading comprehension and vocabulary development through reading. The students were assured that their course credit would not be affected by their participation or performance in the present study.

Before the pre-test, the participants were informed that the target words were not real English words, but they were instructed to infer their meanings as if they were real words. This announcement was not provided in the previous experiments of this study, because the announcement that the target words were pseudowords might change the participants' processing. In this experiment, however, the experiment was conducted over two weeks, and there was a potential danger for the participants to incorrectly learn the pseudowords. Thus, to avoid that situation and make the conditions of the pre- and post-tests equal, the use of pseudowords was informed in advance, not after Session 1.

In the pre-test, the participants were asked to infer the meanings of 16 target words (in boldface) without receiving any prior instruction. The test was paper-based, and eight target words were presented in transparent conditions and the other half in opaque conditions. The time allocation for this test was approximately 12 minutes, based on the results of the first class. After finishing the test, the answer sheet was collected.

Next, the participants completed the lexical inference training. First, to augment the participants' background knowledge of the task, the author explained to them that morphemes of a word are sometimes reliable but sometimes unreliable, and provided examples of transparent (e.g., *sunlight*), opaque (e.g., *honeymoon*), and partially transparent words (e.g., *strawberry*). Then, they were presented with a booklet that included (a) written instructions for the training, (b) sample responses, and (c) the lexical inference worksheet.

Using this booklet, the author instructed the participants on how they could infer unknown words using each of the four steps. For the participants to have a clear understanding of the content of each step and the training process, the author demonstrated how to infer target words for each of the four possible cases: (a) transparent words finishing in Step 2 (*I opened the **mailbox** to check today's newspaper*), (b) opaque words finishing in Step 3 using one morpheme (i.e., *I used a **staircase** to go to the 4th floor*), (c) opaque words finishing in Step 3 using morpheme association (i.e., *The man killed a lot of people and spent 10 years as a **jailbird***), and (d) opaque (deceptively transparent) words finishing in Step 4 (i.e., *I was so tired that I slept on the **carpet** that night*). These instructions lasted approximately 10 minutes. The demonstration was conducted using PowerPoint (Microsoft, 2013), and the participants referred to the presentation slides and sample responses in the booklets. After the demonstration, the author conducted a question-answer session and responded to the participants' questions.

Then the participants completed the lexical inference training, in which they were asked to infer 16 target words at their own pace following the lexical inferencing procedure displayed on the worksheet. The time allocation of the inference training was approximately 25 minutes, which was considered sufficient to complete the task based on observation of the first class. After the booklet was collected, the author announced that they would perform a similar task the following week, but the specific content of the task was not explained.

Session 2 was conducted one week after Session 1 for each of the three classes. Session 2 consisted of the post-lexical inference test (post-test) and the reading proficiency test. The target word-context sentence pairs used for the post-test were identical to those used for each participant in the pre-test. Only their order of presentation was altered from the pre-test. The participants were asked to infer 16 target words for 12 minutes, remembering the inferencing procedure that they had practiced in the inference training. After that, they answered the questionnaire for three minutes, and finally took the 20-minute reading proficiency test.

4.2.2.4 Scoring and analysis

Inference Outcomes

The participants' answers to the pre- and post-tests and in the inference training were categorized using the same criteria as Experiments 1–3. They were classified into *Morphology-Based Interpretations* (MBIs), *Partial Morphology-Based Interpretations* (PMIs), *Context-Based Interpretations* (CBIs), *Inappropriate Interpretations* (INIs), and *None* (see also Appendix 2). MBIs, PMIs, and CBIs were semantically appropriate for transparent words, while PMIs and CBIs were considered appropriate for opaque words. Two raters independently scored 20% each on the pre- and post-tests as well as inference training, with a 91.1% agreement rate in total. Any discrepancies were resolved through discussion, and the remaining data was scored by the author.

To examine whether inference performance improved from (a) the pre-test to the inference training (RQ5-1), and from the pre-test to post-test (RQ5-2), a 3 (time: pre-test, inference training, post-test) × 3 (proficiency: Upper, Middle, Lower) two-way mixed ANOVA was conducted on the proportions of appropriate interpretations for opaque words. In addition, to closely examine qualitative changes, a 3 (time: pre-test, inference training, post-test) × 3 (proficiency: Upper, Middle, Lower) × 5 (type: MBI, PMI, CBI, INI, None) three-way mixed ANOVA was run. The two analyses did not target transparent words, since the participants made appropriate interpretations in most cases, as in Experiment 1–3. Proficiency was treated as a between-participant variable and the others were all within-participant variables.

Semantic Consistency Judgment

The performance of the consistency judgment in Step 2 of the inference training was analyzed. Again, it required the participants to judge whether literal meanings of the two morphemes of target words fit to context semantically. For transparent target words, whose

literal interpretations were consistent with contextual meanings, “Yes” was deemed correct, while “No” was correct for opaque words, as contextual meanings did not allow morpheme-based interpretation to occur semantically. The correct response rates were separately calculated for both transparent words and opaque words.

To examine whether (a) there was any performance difference according to semantic transparency and the participants’ proficiency level, and (b) the correct answer rates were significantly higher than the chance rates (50%), a 3 (proficiency: Upper, Middle, Lower) × 3 (type: transparent, opaque, chance rate) two-way mixed ANOVA was performed.

When interpreting learner performance in this task, the effects of incorrect responses should differ according to semantic transparency. For transparent words, even if the participants thought literal interpretations (i.e., MBIs) did not fit the context, they could still make contextually appropriate interpretations, that is, PMIs and CBIs. It was also observed that they still made MBIs after Step 2. These three types of interpretations were considered appropriate; thus, choosing “No” for transparent words did not always lead to inappropriate interpretations at that stage. In contrast, however, if they perceived that literal interpretations of the morphemes of opaque words were consistent with their surrounding contexts, the judgments directly led to inappropriate interpretations, as they had to finish the inferential process when they answered “Yes” in Step 2. It should also be noted, however, that there were also cases in which participants’ literal interpretations were not morpheme-based.

In addition, examination of both inference outcomes and the semantic consistency judgment will allow us to identify at which step in the procedure the participants had difficulty inferring the meaning of opaque words. For example, if they failed to judge the in/consistency correctly, the failure was attributed to their lack of monitoring the two sources. In contrast, if they correctly judged the target word in Step 2 but ended up with inappropriate interpretations, it was attributed to their meaning-generation skills.

4.2.3 Results

4.2.3.1 Lexical inference outcomes

Table 4.9 presents the descriptive statistics of the lexical inference outcomes in the pre- and post-lexical inference tests, and the training. As in Study 1, while the participants made appropriate interpretations for transparent words in most cases, the proportions for opaque words were much lower, irrespective of time (see Figure 4.9).

Table 4.9

Proportions of Appropriate Interpretations in the Pre-Test, Lexical Inference Training, and Post-Test in Experiment 5

		Pre-test				Inference training				Post-test			
		Transparent		Opaque		Transparent		Opaque		Transparent		Opaque	
Proficiency	*N*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Upper	41	.95	.08	.28	.25	.96	.07	.43	.25	.98	.05	.32	.23
Middle	41	.81	.24	.13	.20	.88	.18	.21	.23	.90	.19	.17	.18
Lower	37	.77	.22	.11	.14	.86	.16	.13	20	.86	.18	.14	.15
Total	119	.84	.21	.17	.21	.90	.15	.25	.26	.91	.16	.21	.20

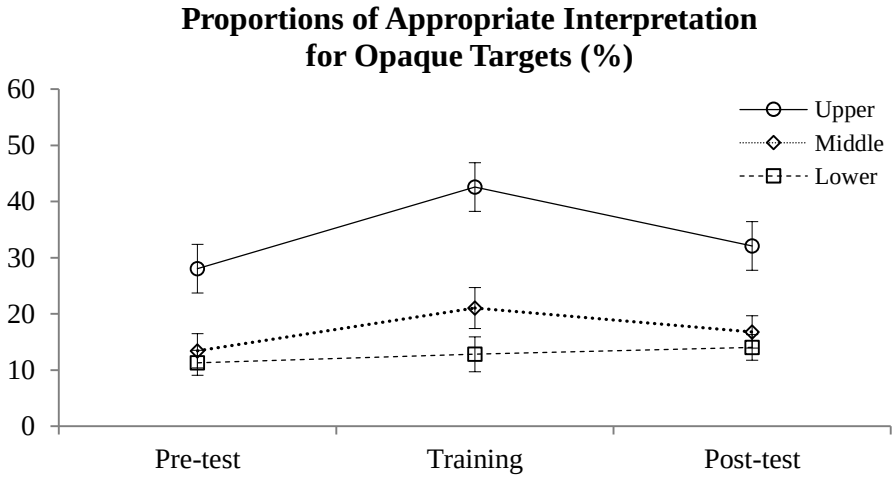


Figure 4.9. Proportions of appropriate interpretations for opaque targets of pre- and post-lexical inference tests and inference training as a function of learner proficiency in Experiment 5.

A 3 (time) × 3 (proficiency) two-way mixed ANOVA was conducted on the proportions of appropriate interpretations of opaque target words (see Table 4.10). The results showed that the main effects of proficiency, $F(2, 116) = 18.69$, $p < .001$, $\eta_p^2 = .238$) and time, $F(2, 116) = 8.41$, $p = .001$, $\eta_p^2 = .068$, were both statistically significant. In addition, the Proficiency × Time interaction was also significant, $F(2, 116) = 2.31$ $p = .042$, $\eta_p^2 = .038$.

Table 4.10

Summary Table for Two-Way ANOVA of the Effects of Time and Proficiency on the Proportions of Appropriate Interpretations for Opaque Words in Experiment 5

Source	*SS*	*df*	*MS*	*F*	*p*	η_p^2
			Within-participants			
Time	0.37	2	0.19	8.41	.001	.068
Time × Proficiency	0.20	4	0.05	2.27	.042	.038
Error (Time)	5.16	232	0.02			
			Between-participants			
Proficiency	3.01	2	1.50	18.11	< .001	.238
Error (Proficiency)	9.63	116	0.08			

To interpret the interaction, a follow-up test was performed. The results showed that the simple main effects of proficiency were significant for all times (all *p*s < .001). Multiple comparisons showed that the Upper group outperformed both the Lower and Middle groups at all times (*p*s < .05). As for the difference between the Middle and Lower groups, the Middle group made more appropriate interpretations than the Lower group only in the inference training, although the difference was only marginally significant ($p = .070$). On the other hand, there was no significant difference in their performance on both the pre- ($p = .638$) and post-test ($p = .549$).

As for the performance difference across time, the simple main effect was significant for the Upper group ($p < .001$), and marginally significant for the Middle group ($p = .076$). On the other hand, the Lower group's performance did not differ over time ($p = .704$). Accordingly, multiple comparisons were conducted for the two groups. In the Upper group, appropriate interpretations increased from the pre-test to the training ($p = .021$), while they decreased from the training to the post-test ($p = .003$). In addition, the difference between the pre- and post-test was insignificant ($p = .237$). Similarly, the performance of the Middle group increased from the pre-test to the training ($p = .021$). In contrast, no significant difference was found between the pre- and post-test ($p = .304$) or between the training and post-test ($p = .198$).

In sum, although the Upper ($M_{diff} = 15.3\%$) and Middle ($M_{diff} = 8.3\%$) groups did better in the inference training than on the pre-test, no groups showed any significant improvement from the pre- to post-test.

However, it is possible that some qualitative changes occurred in their lexical interpretation processes. Table 4.11 provides the proportion of each interpretation type (see also Figure 4.10). To explore qualitative changes in the participants' performances, a 3 (time) × 5 (type) × 3 (Ppoficiency) mixed three-way ANOVA was conducted on the proportion of each interpretation type.

Table 4.11

Proportions of Each Interpretation Type (%) for the Pre- and Post-Lexical Inference Test and Inference Training in Experiment 5

	Transparent					Opaque				
	MBI	PMI	CBI	INI	None	MBI	PMI	CBI	INI	None
					Pre-lexical inference test					
Upper	79.73 (17.52)	12.84 (9.99)	2.70 (6.68)	2.36 (4.92)	2.36 (4.96)	43.92 (26.45)	15.54 (14.56)	12.50 (17.43)	8.78 (11.74)	19.26 (23.50)
Middle	72.26 (24.77)	7.62 (9.63)	1.52 (4.14)	3.05 (6.11)	15.55 (21.25)	51.22 (23.18)	9.76 (12.96)	3.66 (10.17)	7.62 (11.82)	27.74 (23.14)
Lower	67.68 (22.70)	8.23 (8.66)	0.91 (4.32)	3.66 (6.40)	19.51 (23.06)	46.34 (26.85)	7.32 (10.45)	3.96 (8.59)	9.45 (15.76)	32.93 (30.84)
Total	73.00 (22.37)	9.45 (9.62)	1.68 (5.13)	3.05 (5.86)	12.82 (19.90)	42.27 (25.48)	10.71 (13.05)	6.51 (13.00)	8.61 (13.20)	26.89 (26.52)
					Inference training					
Upper	81.76 (15.75)	14.19 (16.18)	0.68 (2.87)	2.70 (5.22)	0.68 (4.11)	19.59 (19.44)	10.81 (9.84)	30.74 (19.45)	30.74 (17.08)	8.11 (14.20)
Middle	77.44 (16.59)	10.06 (11.26)	0.61 (2.73)	3.96 (8.12)	7.93 (16.95)	34.45 (21.06)	6.10 (10.87)	14.33 (17.36)	26.22 (22.33)	18.90 (20.18)
Lower	73.48 (20.19)	12.20 (12.34)	0.61 (2.73)	6.40 (9.32)	7.32 (12.49)	36.59 (20.80)	3.96 (7.09)	7.93 (16.24)	32.93 (21.97)	18.60 (18.57)
Total	77.42 (17.85)	12.08 (13.31)	0.63 (2.75)	4.41 (7.90)	5.46 (12.88)	30.57 (21.63)	6.83 (9.73)	17.23 (19.93)	29.94 (20.73)	15.44 (18.46)
					Post-lexical inference test					
Upper	82.77 (10.77)	13.18 (10.60)	2.03 (5.52)	0.00 (0.00)	2.03 (4.67)	53.72 (25.66)	19.26 (15.75)	12.84 (13.01)	6.08 (8.13)	8.11 (13.57)
Middle	77.44 (21.14)	10.98 (10.53)	1.22 (3.76)	1.52 (4.14)	8.84 (17.51)	62.50 (27.10)	11.89 (13.68)	4.88 (8.78)	7.93 (13.04)	12.80 (19.06)
Lower	73.78 (20.12)	10.67 (10.67)	1.83 (4.47)	2.13 (6.19)	11.59 (18.41)	57.01 (22.54)	9.76 (10.27)	4.27 (10.31)	8.23 (10.68)	20.73 (26.75)
Total	77.84 (18.37)	11.55 (10.57)	1.68 (4.58)	1.26 (4.43)	7.67 (15.53)	57.88 (25.20)	13.45 (13.82)	7.14 (11.34)	7.46 (10.83)	14.08 (21.19)

Note. Standard deviations are in parentheses.

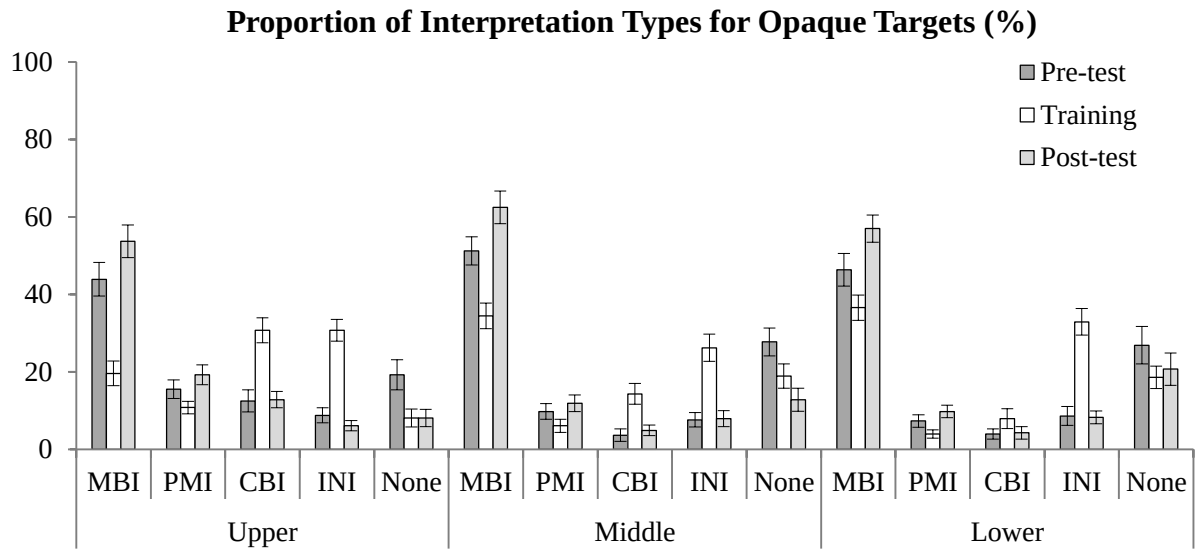


Figure 4.10. Proportion of each interpretation type for opaque target words in Experiment 5.

The results showed that main effects of time, type, and proficiency were all statistically significant (all *p*s < .05). In addition, the interactions of Type × Proficiency and Time × Type were significant (*p*s < .001), while the Time × Proficiency interaction did not reach a level of statistical significance ($p = .336$). Importantly, the Time × Type × Proficiency three-way interaction was significant, $F(2, 116) = 1.58$, $p = .042$, $\eta_p^2 = .027$. To answer the research questions, post-hoc tests were conducted to analyze how interpretation trends in each proficiency group differed from (a) the pre-test to the inference training and (b) from the pre- to post-test.

First, we will examine the changes from the pre-test to the inference training (RQ5-1). All groups used fewer MBIs ($p < .01$) and the None type (all *p*s < .001), suggesting that the training had the effect of preventing literal interpretations among all participants. As for appropriate interpretations (i.e., PMI, CBI), only the Upper group made fewer PMIs ($p = .011$), and the proportions of CBIs increased in the Upper and Middle groups (both *p*s < .001). This means that the increase in appropriate interpretations in the two groups was due to their increase

in context-based interpretations. However, at the same time, the participants in all groups made more INIs (all *p*s < .001).

Next, we will examine the changes from the pre- to post-test (RQ5-2). Post-hoc tests showed that proportional differences were observed in the two interpretation types regardless of proficiency level: the proportions of MBIs increased, while the None type decreased (all *p*s < .001). This means that the participants tried to infer more unknown words in the post-test than in the pre-test, but they resulted in inappropriate, literal interpretations.

4.2.3.2 Semantic consistency judgment

Table 4.12 shows the descriptive statistics of the correct response rates of consistency judgments in Step 2 in the inference training. Given the chance rate of this task (50%), the average correct answer rates for all the participants (Transparent: 64.3%; Opaque: 66.5%) were considered to be relatively low. In addition, the proportion of opaque words means that they made incorrect literal interpretations for the rest of the cases (35.7%) because they perceived that literal interpretations fit the context, though some participants produced non-literal interpretations in Step 1.

Table 4.12

Descriptive Statistics of Semantic Consistency Judgments in Step 2 of the Inference Training in Experiment 5

		Transparent			Opaque			ALL		
Proficiency	*n*	*M*	95% CI	*SD*	*M*	95% CI	*SD*	*M*	95% CI	*SD*
Upper	37	.63	[.55, .71]	.26	.83	[.77, .89]	.18	.73	[.68, .78]	.15
Middle	41	.65	[.59, .71]	.21	.62	[.53, .71]	.29	.63	[.58, .69]	.18
Lower	41	.65	[.57, .72]	.24	.54	[.46, .62]	.25	.59	[.54, .65]	.17
Total	119	.64	[.60, .68]	.23	.66	[.61, .71]	.28	.65	[.62, .68]	.18

As for transparent words, one may wonder why the correct response rates (Upper: 63.1%; Middle: 65.6%; Lower: 65.2%) were lower than those of MBI in the training (Upper: 81.7%; Middle: 77.4%; Lower: 73.5%). This was because there were many cases in which the participants used MBIs for transparent words even after they chose "No" in the consistency judgment.

A 3 (type) × 3 (proficiency) two-way mixed ANOVA was performed on the correct response rates of the consistency judgment (Table 4.13). The results showed a main effect for proficiency, $F(2, 116) = 6.89$, $p = .001$, $\eta_p^2 = .106$ and type, $F(2, 116) = 24.22$, $p < .001$, $\eta_p^2 = .173$. In addition, the Type × Proficiency interaction was also significant, $F(2, 116) = 7.80$, $p < .001$, $\eta_p^2 = .118$.

Table 4.13

Summary Table for Two-Way ANOVA of the Effects of Type and Proficiency on the Correct Response Rates for Consistency Judgments in Experiment 5

Source	*SS*	*df*	*MS*	*F*	*p*	η_p^2
			Within-participant			
Type	1.93	2	0.97	24.22	< .001	.173
Type × Proficiency	1.24	4	0.31	7.80	< .001	.118
Error	9.25	232	0.04			
			Between-participant			
Proficiency	0.52	2	0.26	6.89	.001	.106
Error	4.37	116	0.04			

A subsequent test showed that the Upper group outperformed the Middle and Lower groups for opaque words (*p*s < .001), while correct response rates did not differ among the groups for transparent words ($p = .910$). In addition, it was found that the Upper and Middle groups made correct responses at a significantly higher rate than chance, irrespective of

transparency (*p*s < .010), while the Lower group's correct response rate for opaque words was lower than the chance rate (p = .348), suggesting that they failed to perceive the semantic consistency of the two sources for opaque words.

4.2.3.3 Questionnaire

Table 4.14 displays the descriptive statistics of the questionnaire conducted after the post-test to examine the participants' perception of the inference training. The average ratings suggest that, although the training was difficult for tehm, the participants thought that they could somewhat follow the inferencing procedure in both the training and post-test.

Table 4.14

Descriptive Statistics of the Questionnaire in Experiment 5

		Q1 (Perceived difficulty of training)			Q2 (Perceived task completion during training)			Q3 (Perceived task completion during post-test)		
Proficiency	*n*	*M*	95% CI	*SD*	*M*	95% CI	*SD*	*M*	95% CI	*SD*
Upper	37	2.75	[2.46, 3.04]	0.87	3.72	[3.49, 3.95]	0.70	3.61	[3.42, 3.81]	0.60
Middle	41	2.49	[2.18, 2.79]	1.00	3.37	[3.10, 3.63]	0.86	3.49	[3.22, 3.75]	0.87
Lower	41	2.10	[1.87, 2.32]	0.74	3.15	[2.81, 3.49]	1.11	3.32	[2.99, 3.65]	1.08
Total	119	2.43	[2.27, 2.60]	0.91	3.40	[3.23, 3.57]	0.93	3.47	[3.31, 3.63]	0.88

Note. Q1 ranged from 1 [*very difficult*] to 5 [*very easy*], Q2 and Q3 ranged from 1 [*not at all*] to 5 [*very well*]).

To see if there was any difference in the participants' perception of the procedure according to proficiency level, a one-way ANOVA was performed on each question. A main effect of proficiency was found for Q1, $F(2, 116) = 5.42$, $p = .006$, $\eta_p^2 = .086$, and Q2, $F(2, 116) = 3.85$, $p = .024$, $\eta_p^2 = .063$, but not for Q3, $F(2, 116) = 1.08$, $p = .342$, $\eta_p^2 = .018$. A post-hoc

test revealed that the Upper and Middle groups rated higher than the Lower group for both Q1 and Q2 (*p*s < .01). These results indicate that perceived difficulty and performance in the training (Session 1) differed according to proficiency level.

4.2.4 Discussion

Does lexical inference training using an inference worksheet support the lexical inference performance of beginner-level Japanese EFL learners? (RQ5-1)

Experiment 5 was designed to examine the effects of lexical inference training for beginner-level EFL learners. For this purpose, the training used an inference worksheet that explained when and how to utilize morphological and contextual information according to the semantic transparency of morphemes in unknown compound words.

To examine whether their inference performance was enhanced by the training, the results of the pre-test and the lexical inference training were compared. Overall, the results of the two-way ANOVA on the proportion of appropriate interpretations of opaque target words showed that the Upper and Middle groups (although only marginally) improved in the inference training from the pre-test (Upper: 28.0% → 42.6%; Middle: 13.4% → 21.0%). In contrast, the performance of the Lower group did not significantly improve (11.3% → 12.8%).

The analysis regarding interpretation type, however, indicated that the training had qualitative effects. The participants in all proficiency groups used significantly fewer MBIs for opaque words in the training (Upper: 43.9% → 19.6%; Middle: 51.2% → 34.5%; Lower: 46.3% → 36.6%). The semantic consistency judgment in Step 2 was designed to help the participants consider the semantic relationship between the two sources and explore context-based meanings, rather than literal interpretations. However, the increase of INIs in the training for all groups (Upper: 8.8% → 30.7%; Middle: 7.6% → 26.2%; Lower: 9.5% → 32.9%) indicates

that generating context-based meanings in Steps 3 and 4 often resulted in inappropriate interpretations in many cases.

As for appropriate interpretations, CBIs were used more often by the Upper (12.5% → 30.7%) and Middle (3.7% → 14.3%) groups in the training, which accounted for the increase in appropriate interpretations. This means that the worksheet made them consider the context-based meaning after perception of semantic inconsistency. In contrast, PMIs were used less often by the Upper group (15.5% → 10.8%), while no difference was seen in the Middle and Lower groups: PMIs were rarely used in these two groups in both the pre-test and training (Middle: 9.8% → 6.1%; Lower: 7.3% → 4.0%). These results suggest the difficulty of integrating the semantic information of morphemes and context, like PMI. In the training, the participants were encouraged to consider contextually appropriate interpretations that included morphological information in Step 3 before making fully context-based inferences in Step 4. The difficulty of integrating the two sources of information in lexical inference-making has also been found in previous studies, even in multiple-choice format (Hamada, 2014). However, given the fact that they were explicitly asked to consider the integrated meanings in the training, the decrease in PMI use in the Upper group might reflect their perceived difficulty of the integration of semantic information from the two sources and their preference for CBIs.

The overall results indicate that the effectiveness of the lexical inference training in Experiment 5 can be interpreted as rather small given that the inference worksheet explicitly showed participants when and how they should use morphemes and context for inferences in each step. Therefore, the following discussion focuses on why the participants did not benefit as much from the training as expected by examining their performance in consistency judgments and patterns in their interpretations.

First, the correct answer rates of the semantic consistency judgment in Step 2 showed that the participants often made incorrect responses in both transparent (Upper: 36.9%; Middle:

34.4%; Lower: 34.8%) and opaque words (Upper: 16.8%; Middle: 37.5%; Lower: 45.1%). In this regard, the proportions of transparent words were not significant because there was a chance that they subsequently made contextually appropriate interpretations (i.e., MBI, PMI, CBI). However, the high proportions of incorrect judgments for transparent words in all groups indicates that the more accurate understanding of context was required for transparent words, as suggested by Experiment 4.

For opaque words, incorrect responses directly led to inappropriate literal interpretations, though some literal interpretations produced in Step 1 were not literal ones. Higher proportions of participants in the Middle and Lower groups relative to the Upper group showed that they often failed to understand the semantic consistency between morpheme and context in Step 2, even when their attention was directed to both sources of information. The correct response rate of the Lower group (55.0%) was not statistically higher than the chance rate (50%), indicating that they had particular difficulties in understanding the semantic inconsistency between the two sources, let alone the subsequent required meaning generation.

However, the lack of monitoring of the two sources did not fully explain their performance because the proportions of appropriate interpretations were much lower than the correct responses of consistency judgments. This means that they often made inappropriate inferences even if they could perceive the semantic inconsistency. In the training, all participants had more INIs than in the pre-test (Upper: 30.7%; Middle: 26.2%; Lower: 33.0%).

This gap between the correct answer rates of the consistency judgment and the proportion of appropriate interpretations raises a question. Why did the participants, who successfully perceived the inconsistency, subsequently fail to make appropriate interpretations? This is because the large proportions of INIs were also seen in the Upper group, who correctly judged the consistency in most cases (83.2%). Furthermore, in this study, successful consistency judgments should be almost synonymous with successful inferences due to the following

features. First, the present materials consisted of simple vocabulary and sentence structure; thus, once they could understand the context, the act of generating contextually-based meanings itself was not considered demanding for university-level participants. In addition, in Steps 3 and 4, the participants were instructed to write down the meanings they could come up with only if they were consistent with their contexts. Accordingly, the scoring criteria also regarded any answers that would be contextually plausible as appropriate. Given these features of the materials, task instruction, and scoring criteria, understanding inconsistency should be largely synonymous with their success in generating appropriate interpretations.

A possible reason for this is the different degree of context understanding required for the two tasks. That is, making appropriate interpretations requires more precise text understanding than perceiving the consistency of the two sources of information. Making appropriate interpretations requires readers to narrow down the semantic meaning based on the surrounding context; thus, it requires accurate understanding of the context sentence. On the other hand, perception of inconsistencies can be achieved only if readers understand more superficial information, such as the topic of the context sentence and the morpheme-based meanings of the target word. Thus, an accurate understanding of the context is not necessarily required to correctly judge semantic in/consistency. Accordingly, the participants made semantic consistency judgments based on imperfect understandings of the context sentences, which led to their failure to generate appropriate interpretations.

Does lexical inference training enhance the lexical inference skills of beginner-level Japanese EFL learners? (RQ5-2)

It was found that the lexical inference training had some positive effects on beginner-level learners' inferencing skills for opaque words, such as (a) avoidance of literal interpretations and explorations of context-based meanings (all groups) and (b) slight increase

of appropriate interpretations (the Upper and Middle groups). This section discusses whether these effects were retained one week after the training in the post-test, where the worksheet was not provided. For this purpose, the participants' performance on the pre- and post-lexical inference tests were compared.

The results showed that learner performance did not differ in terms of either proportions of appropriate interpretations or types of interpretations. These results suggest that the positive effects of the inference training were not retained in the post-test. In terms of inference outcomes, these results suggest that the participants behaved similarly in both lexical inference tests before and after the training.

There are two possible reasons that could explain the null effects of the training. The first is that the participants may have forgotten the steps of the inferencing procedure. Although they practiced the inferencing procedure for about 25 minutes and inferred 16 unknown words, a one-shot training model may be insufficient for learners to internalize the process. On the one hand, it seems quite natural because strategy training requires a lot of time to be effective (Fraser, 1999). On the other hand, it can be argued that the effects of the inference training in the present study appeared immediately during training in that they could follow the procedure, aside from the accuracy of the performance, also supported by their response for Q2 of the questionnaire regarding their perceived task completion during the training (M = 3.40). This is possibly because the training focused on how to infer unknown words based on their semantic transparency, rather than a general lexical processing strategy. Therefore, it would be unlikely that they forgot the contents of the training completely. In addition, their answers to Q3 of the questionnaire indicated that they thought they could somewhat complete the task in the post-test based on their memory of the inferencing procedure (M = 3.47). Although these data are from subjective ratings, they suggest that the participants tried to replicate the procedure in the post-test.

Another possibility is that they remembered the content of the procedure to some extent, but could not replicate it in the post-test. This is related to the difference in the cognitive complexity of the inference training and the post-test. In the training, the participants could follow the inferencing procedure one step at a time while referring to the worksheet. However, in the post-test, they were only asked to produce the meanings of words independently by recalling the procedure they had practiced in the previous week. Accordingly, they had to generate the final products (inferred word meanings) by carrying out multiple cognitive processes. This explanation is also supported by the participants' answers to Q1 of the questionnaire, in which they responded that the inference training was rather difficult for them (M = 2.43). This means that following the procedure was a cognitively demanding task, even when the inferencing procedure worksheet was available. Accordingly, completing the multiple cognitive processes was considered much more difficult for the less proficient learners, who had limited processing skills. The cognitive burden of each step likely left few cognitive resources to organize their overall lexical inferencing process.

4.2.5 Conclusion of Experiment 5

The aim of Experiment 5 was to (a) examine the effects of inference training for beginner-level EFL learners and to (b) identify the specific source of difficulties for learners in the process of inferring opaque unknown words, to provide them with the necessary support. In the training, the beginner-level participants, who were further divided into three groups, inferred unknown words using a worksheet that described when and how to use linguistic clues.

The results showed that this training had limited beneficial effects. It helped them avoid making literal interpretations for opaque words and contextually appropriate meanings; however, it led to a only slight increase of appropriate interpretations for the Upper and Middle groups, while no significant increase was observed for the Lower group. Furthermore, improved

performance between the pre- and post-tests was not observed in all groups, suggesting that the above effects of the training were not observed one week later.

The error analysis identified two types of difficulties. For the Middle and Lower groups, especially the latter group, the learners had difficulties understanding the semantic relationships between morphemes and their contexts, which was demonstrated by the lower correct response rates for the consistency judgments. In addition, the higher proportions of inappropriate interpretations in all groups showed that they also had difficulty generating contextually appropriate meanings even when they were asked to do so. The possible cause of the two processes was attributed to their inaccurate understanding of the context. In addition, following the inferencing procedure used in the training was difficult for the participants without a worksheet due to insufficient cognitive and linguistic skills. Therefore, the above-mentioned training benefits were not retained one week after the training.

4.3 Conclusion of Study 2

Study 2 focused on the lexical inferencing skills of beginner-level EFL learners. It has frequently been reported that beginner-level EFL learners often make incorrect inferences because of their overreliance on morphemes without considering the contextual meaning because of their limited linguistic skills (Hamada, 2014; Huckin & Coady, 1999). However, the lack of detailed investigations of such learners makes it difficult to know which lexical inferencing processes are problematic for them, and accordingly, how to support them in their learning. Accordingly, Experiments 4 and 5 were conducted to identify the processes in which beginner-level learners experience difficulties and examine the effects of interventions for such learners.

Experiment 4 examined the participants' understanding of the semantic relationship between morphological and contextual information using on- and off-line semantic consistency

judgment tasks. The results revealed that the participants generally understood the semantic relationship between the two sources, possibly due to the explicit direction of their attention to the relation. These findings led us to the tentative conclusion that their difficulty lies in meaning-generation processes.

Based on the findings of Experiment 4, Experiment 5 examined the effects of lexical inference training focusing on how to generate appropriate target word meanings. The participants practiced inferring unknown words using a worksheet that described how to use morphological and contextual information according to the semantic transparency of morphemes in a step-by-step manner. The results revealed that the use of the worksheet helped them to consider both sources of information and explore contextually appropriate meanings. However, the participants' insufficient contextual understanding prevented them from understanding the semantic relationship between the two sources, and subsequently from generating contextually appropriate inferences. Furthermore, the inferencing process they practiced during the training was not applied in the post-test in their normal lexical inferencing.

Overall, the results of Study 2 indicate that beginner-level learners had difficulties in both their perception of the semantic relationship between the two sources of information and meaning-generation processes depending on their proficiency level. However, failure in both processes was caused by their lack of accurate understanding of contextual information, which could, in turn, hinder their effective use of the inferencing procedure. These findings suggest that effective use of morphological and contextual information during lexical inferencing is a much more demanding task for beginner-level learners and requires more support than it does for intermediate-level learners.

Chapter 5

General Discussion

In order to examine the processing and interpretation of unknown words with morphological and contextual information among Japanese EFL learners, the current research conducted two studies that comprised of five experiments (Experiments 1–5). The two experimental studies targeted learners belonging to different proficiency bands: Study 1 focused on intermediate-level learners (Experiments 1–3) and Study 2 examined beginner-level learners (Experiments 4 and 5). This chapter overviews the findings and discusses them generally. As mentioned in Section 2.6, the unresolved issues differed according to the learner proficiency; however, the use of a common framework involved in the interpretation process will allow us to characterize and compare the findings from the two groups. Appendix 6 summarizes the English proficiency levels of the participants of Experiments 1–5.

5.1 Processing and Interpreting Unknown Words and Textual Information Based on Morphological and Contextual Information of Intermediate-Level Learners

The three experiments in Study 1 targeted intermediate-level Japanese EFL learners (Experiments 1–3) to examine their processing and interpretations of unknown words during normal reading and lexical inferencing. In addition, their resulting text representations were also examined. The eight research questions addressed in Study 1 are as follows.

RQ1-1: Are intermediate-level Japanese EFL readers sensitive to semantic relationship between morphological and contextual information during reading?

RQ1-2: What types of information do intermediate-level Japanese EFL readers use to interpret transparent and opaque unknown words?

RQ 1-3: What types of information do intermediate-level Japanese EFL readers revise to achieve coherent representations of a sentence with opaque unknown words?

RQ 2-1: How do intermediate-level Japanese EFL readers process passages with transparent and opaque unknown words?

RQ 2-2: What kinds of semantic representations do intermediate-level Japanese EFL readers construct for unknown words through reading?

RQ 3-1: How well do intermediate-level Japanese EFL learners infer unknown words in a lexical inferencing task according to the semantic transparency of the morphemes?

RQ 3-2: What kinds of inferential strategies are related to the appropriate interpretations for opaque unknown words?

RQ 3-3: What are the causes of inappropriate word-based interpretations of for opaque words?

First, Experiment 1 examined whether intermediate-level Japanese EFL learners were sensitive to the semantic relationship between morphemes and the context during reading (RQ1-1), and interpret unknown words according to the semantic transparency of unknown words' morphemes (RQ1-2). In addition, their sentence representations were examined (RQ1-3). The participants first read the sentential context with either transparent or opaque unknown words with a word-by-word self-paced reading task, and then translated the whole context into Japanese. The results indicated that while they were able to perceive the semantic relations between the two sources, both in informative and neutral contexts, they made literal interpretations of opaque words in about half the cases, suggesting the difficulty of interpreting opaque unknown words. In addition, they sometimes changed or distorted the interpretations of the context so that literal interpretations would make sense within the context.

In Experiment 1, the participants' interpretations of target words were examined through

a translation task; thus, it was unclear what kinds of semantic representations they constructed during reading. In addition, it was expected that single-sentence context might be insufficient compared with normal texts, where more contextual information is available.

Thus, Experiment 2 examined the processing of passages with either transparent or opaque words (RQ2-1) and its relation with semantic memory representations of unknown words (RQ2-2). To this end, the experiment adopted two-sentence passages (read sentence by sentence), and their interpretations of target words were assessed via comprehension questions presented immediately after the passage reading. The reading time suggested that they were sensitive to the semantic relationship between the two sources, and would have tried to infer words' meanings even after sentences containing unknown words (second sentences); they employed varying degrees of mental effort by the context quality of the first sentences. However, this attempt did not always lead to contextually appropriate interpretations: they often interpreted unknown opaque words literally or did not provide any answers.

The apparent discrepancy between the perceptions of the relations between morphemes and the context and subsequent interpretations found in the two experiments, as well as the large individual differences in the interpretation trend in Experiment 2, raised the possibility that lack of focus on unknown words during reading solicited insufficient mental effort to infer the unknown words, resulting in high proportions of literal interpretations.

Experiment 3, therefore, examined the participants' interpretations in a lexical inferencing task to confirm the effects of affective importance of unknown words (RQ3-1). The experiment adopted the think-aloud method to elucidate the inferential processes associated with appropriate interpretations (RQ3-2) and the causes of inappropriate literal interpretations (RQ3-3). The results showed that EFL learners generally made contextually appropriate interpretations of opaque words, unlike the previous experiments. These interpretations were associated with the use of metacognitive strategy induced by semantic conflicts, and linguistic

clue that used a part of morphological information and fully context-based inferences, rather than literal interpretations of morphemes. In addition, the qualitative analysis revealed that literal interpretations of opaque words were sometimes made by changing the interpretation of the context, indicating their mental effort to achieve meaningful interpretations.

Based on the summary of the results of Experiment 1–3, we will discuss two perspectives: (a) the causes of inappropriate interpretations of opaque words and related factors, and (b) the representations of text with opaque unknown words.

The causes of inappropriate interpretations of opaque words and conditions in which appropriate interpretations are achieved

Table 5.1 presents the proportions of each interpretation type for opaque words in Experiments 1–3. Similar to past studies (Bensoussan & Laufer, 1984; Huckin & Bloch, 1993; Laufer, 1989; Mori, 2002), the participants of this study often made such literal interpretations (i.e., MBI), especially in Experiments 1 and 2. At the same time, they made appropriate interpretations in some conditions. Therefore, the discussion here focuses on the difficulties in appropriate interpretations, and factors affecting the performance.

Table 5.1

Proportions of Each Interpretation Type (%) for Opaque Words in Experiments 1–3

	Informative context					Neutral context				
Experiment	MBI	PMI	CBI	INI	None	MBI	PMI	CBI	INI	None
Experiment 1	48.33	22.08	23.33	2.50	1.25	46.67	35.00	15.00	0.42	2.08
	(28.27)	(16.38)	(17.30)	(6.77)	(3.75)	(23.92)	(18.09)	(13.84)	(2.24)	(4.66)
Experiment 2	29.17	13.69	28.65	1.64	26.94	40.92	13.54	19.49	1.04	25.00
	(31.84)	(13.75)	(21.26)	(4.43)	(30.71)	(32.31)	(18.77)	(17.94)	(3.53)	(32.13)
Experiment 3	7.29	37.58	49.65	2.90	2.31	11.46	60.71	13.84	7.76	6.23
	(11.56)	(22.08)	(24.92)	(8.21)	(6.16)	(8.76)	(21.90)	(17.76)	(10.25)	(8.09)

Note. Standard deviations are in parentheses.

Findings from past studies have attributed readers' misidentification of unknown words to their misapprehension about words (Bensoussan & Laufer, 1984; Huckin & Bloch, 1993; Laufer, 1989) and insufficient attention to context and monitoring of the two sources (e.g., Hamada, 2014). However, the present findings offer another interpretation of such literal interpretations.

In Experiments 1 and 2, the participants often made literal interpretations of opaque words. However, the longer reading times for opaque conditions in the reading task suggest that they had noticed the semantic conflicts conveyed by the two sources. It is noteworthy that, compared to normal reading and sentence-by-sentence reading in Experiment 2, the reading style in Experiment 1 was somewhat more cognitively demanding than normal reading due to the word-by-word presentation, where they were unable to read back to the preceding portion of the sentence. Given that the detection took place upon encountering the target words in Experiment 1 demonstrate that they made literal interpretation even after knowing or noticing that literal interpretations were not semantically appropriate.

In addition, such longer reading times for opaque conditions were observed not only for informative context (Experiment 1 [target word reading time]: M_{diff} = 208.0ms; Experiment 2 [reading time of first sentence/syllables]: M_{diff} = 264.1ms) but also for neutral contexts (Experiment 1: M_{diff} = 179.1ms; Experiment 2: M_{diff} = 266.2ms). These results indicate that the perception of semantic relation would not be affected by how much information is derived from the surrounding context, only if the semantic information conveyed by the two sources were incongruent.

The readers' strong sensitivity, however, may seem inconsistent with the above-mentioned causes of the misidentifications. This discrepancy can be attributed to the sensitive measures employed in this study. The discussion in past studies was mostly based on the participants' verbal reports in a think-aloud procedure (e.g., Huckin & Bloch, 1993) and

assumptions based on lexical inferencing outcomes (Hamada, 2014; Laufer, 1989). Although think-aloud comments would reflect readers' real time processing to a certain degree, this method would not be sensitive enough to capture the process of detecting the inconsistency. In this regard, the results of Experiment 3, which also used the think-aloud method, support this possibility. Although successful interpretations of non-literal expressions involve the participants' initial awareness of the inconsistency (Cacciari & Levorato, 1995, 1999; Oakhill et al., 2016), proportions of Stating difficulty or failure (Informative: 24.7%; Neutral: 20.1%) and Questioning (Informative: 11.5%; Neutral: 20.8%) were much lower than the proportions of appropriate interpretations (Informative: 87.2%; Neutral: 74.6%). Since most of these strategies were often made for the response to the semantic inconsistency in this study, the smaller proportions suggest that such cognitive processes are less likely to be reflected in think-aloud comments. Accordingly, the use of sensitive measures allowed us to suggest that the readers' incorrect word-based interpretations, including misidentifications of deceptively transparent words, are not necessarily a result of their failure to consider contextual information, though their perceptions of the potential inconsistencies were unconscious ones.

However, one may still argue that such slight response might be insignificant unless the detection contributes to appropriate interpretations. In reference to this, reading time data in Experiment 2 suggested that the participants try to infer the meanings of target words even in the second sentence, followed by the initial detection of the inconsistency in the first sentence. Similarly, an idiom study showed positive correlations between longer reading times for figurative expressions relative to literal counterparts and subsequent correct interpretations of those figurative expressions (Oakhill et al., 2016). Thus, sensitivity to the relationship between morphological and contextual information would be an important first step to achieve correct interpretations of opaque words.

Therefore, the source of difficulty in interpreting opaque words lies in generating

contextually appropriate meanings after noticing the semantic inconsistency between the two sources. However, these results do not mean that intermediate-level learners are unable to infer opaque words. This is because these trends were observed in Experiments 1 and 2, which examined their interpretation via translation protocols and comprehension questions, respectively, but not in Experiment 3. The better performance in Experiment 3 would be related to the importance of the words in the lexical inference task.

However, this does not mean that EFL readers, in normal reading or in a translation task, do not always sufficiently engage in inferential processes due to the significant distribution patterns of MBI, PMI, and CBI in Experiments 1 and 2, suggesting the effects of individual differences between the participants. The difference can be attributed to the individual perceptions of unknown words during reading; some might have thought it was not necessary to infer the word meanings while others might have thought they should infer the concrete meanings of words, suppressing literal interpretations. This was also supported by the high proportions of None (*I don't know*) for opaque words in Experiment 2, implying that they were somewhat certain about the inappropriateness of literal interpretations, and that they chose not to provide any answers rather than try to generate contextually appropriate meanings. Collectively, whether readers make appropriate interpretations of opaque words can be ultimately determined by their perceived importance of the words, not necessarily a task requirement.

It should be noted that when learners were determined to infer the meaning, how the learners achieve appropriate interpretations (i.e., PMI, CBI) can be affected by the quality of the context surrounding unknown words. In Experiments 1 and 3, in which a single sentence context was used, the participants made more frequent PMIs in neutral contexts (Experiment 1: 35.0%; Experiment 3: 60.7%) than in informative ones (Experiment 1: 22.0%; Experiment 3: 37.6%). On the other hand, the proportions of CBI showed the opposite trend; it was made more

frequently in informative contexts (Experiment 1: 23.3%; Experiment 3: 49.6%) than in neutral ones (Experiment 1: 15.0%; Experiment 3: 13.8%). This indicates that a relative amount of context information affects their inclusion of morpheme-based meanings, i.e., PMI and CBI. In neutral context, lack of sufficient semantic information from context made them include morpheme-based meaning, while concrete semantics derived from informative contexts induced them to make interpretations free from morphological information.

The results of Experiment 2, which adopted a two-sentence context (neutral or informative first sentence and directive second context), further indicated the effects of context on types of interpretations. In Experiment 2, while the proportions of PMI did not differ with context quality of the first sentence (informative: 13.7%; Neutral: 13.5%), those of CBI were more frequent in informative conditions (28.7%) than neutral ones (19.4%). This means that the addition of subsequent directive context would have reduced the effects of the context quality of the first sentence; hence, the proportion of PMI was not affected. This suggests how EFL learners achieve appropriate interpretations of opaque words can be affected to a great extent by the contextual information surrounding them. The result is consistent with the result of qualitative analysis of Ushiro et al. (2013), which showed that EFL learners made fewer incorrect primary-meaning based inferences for informative contexts.

Finally, based on the discussion, I propose a framework regarding how intermediate-level EFL learners interpret opaque unknown words, along with two key processes, (a) perceptions of semantic inconsistency between the two sources and (b) generation of contextually appropriate meanings, based on the basic assumptions of GEM (cf. Oakhill et al., 2016). Since the two processes are common with the discussion of beginner-level learners, introduced in the next section, this framework serves to highlight the role of learner proficiency. In addition, to capture the comprehensive understanding of the relationship, influencing factors of their decision, unique to intermediate-learners, were also incorporated, as illustrated in Figure 5.1.

First, when readers are faced with opaque words, the first turning point is whether they can notice the inconsistency between the morpheme-based and context-based meanings. In this regard, intermediate-level learners were uniformly able to perceive the inconsistency even during reading, regardless of context informativeness.

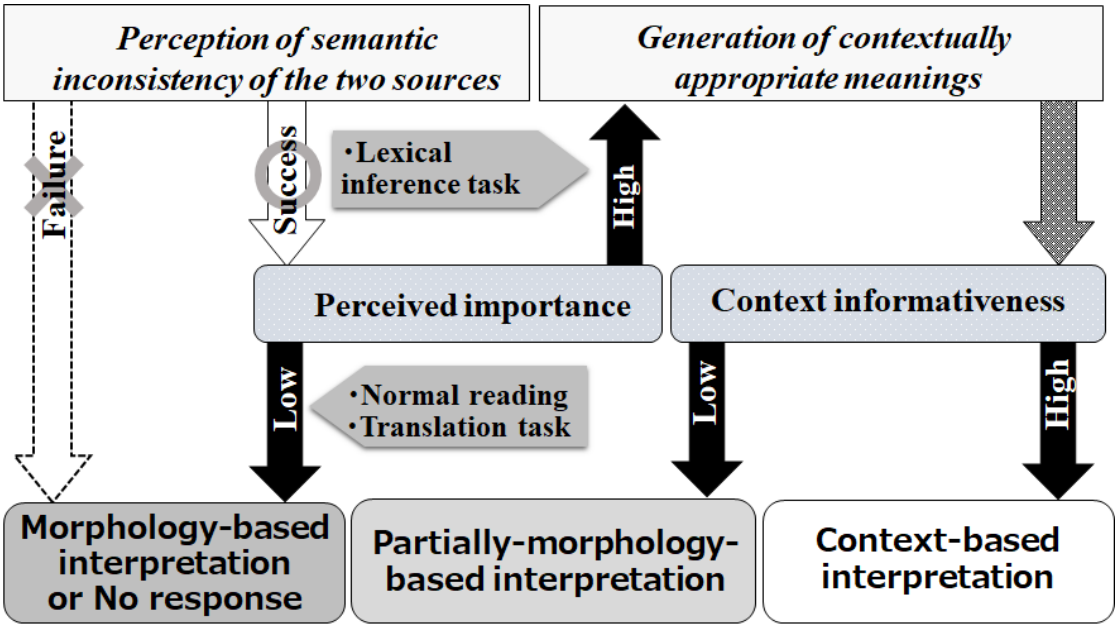


Figure 5.1. Relationship among the processes involved in inferencing opaque words, interpretation types, factors affecting the inferential processes (i.e., perceived importance of unknown words, context informativeness), and tasks employed in this study. The gray circle represents the process was made uniformly, while gray x indicates the process was hardly seen.

Unlike GEM, readers do not always tryt to generate contextually appropriate meanings after noticing the inconsistency. At this time, readers' perceived importance of target words could influence further exploration of contextually appropriate interpretations (high importance), result in literal interpretations (i.e., MBI), or lead to no response (low importance), though their perceived importance is generally high in a lexical inference task.

Finally, when trying to generate contextually appropriate meanings (i.e., PMI, CBI),

context quality could affect their interpretations. When the context contains neutral information for unknown words (low informativeness), readers tend to consider interpretations that include morpheme-based meanings (i.e., PMI), while fully context-based interpretations (i.e., CBI) are likely to be made (high informativeness) as the informativeness increases. In this regard, the context informativeness is determined by the total amount of surrounding context that readers could use, and not restricted to the single sentence, given the difference in the interpretation trend between Experiments 1, 3 (single-sentence context), and Experiment 2 (two-sentence passage).

In conclusion, intermediate-level Japanese EFL learners had the necessary skills and metacognition. Further, their processing and interpretation of opaque unknown words can be affected by motivational and textual factors (i.e., perceived importance of the words, context informativeness).

Representations of a text with opaque unknown words

Some studies have argued that wrong word-based inference leads to global misinterpretations since mistakenly identified words could be the potential clues for inferring other unknown words (Laufer, 1989). In addition, Bensoussan and Laufer (1984) reported the case in which such incorrect inferences made a reader distort the interpretations of subsequent context. Given these potential negative effects of wrong word-based inferences, it was important to examine whether such incorrect inferences should be regarded as the problem of local (i.e., lexical) or global (i.e., sentence, text) reading comprehension.

To this end, the present study examined the sentence representations of the context in which opaque unknown words were embedded in the translation task in Experiment 1. In the analysis, translation protocols, in which meaningful interpretations were made, were extracted and the revised information (part of target word, whole target word, context) was analyzed. The

results indicated that while they revised the morphemes of opaque unknown words in 76.3% of the cases (Part of target word: 45.6%; Whole target word: 30.7%), in the remaining cases, however, they revised the interpretations of context information (Context: 16.3%), and both context and morphemes (Part of target word × Context 6.3%; Whole target word × Context: 1.1%).

These results indicated that they changed or distorted the contextual meanings. In addition, unlike the cases of misidentification, these interpretations of contexts were probably made intentionally given their awareness of the semantic inconsistency during reading. To maintain the coherence of sentence meaning, in some think-aloud comments of Experiment 3, some participants retained the literal interpretations of morphemes and the changed context interpretations after noticing the semantic conflicts (Modification of context, Elaboration of context). Therefore, although such revision of context interpretations reflects their conscious mental effort to maintain coherent representations of text (Ushiro, Mori, et al., 2016), prioritizing morpheme-based meanings over known context is undesirable and leads to comprehension problems.

In addition, the result of Experiment 2 suggested that EFL learners made literal interpretations even after reading directive context that followed the sentence containing the unknown words (Informative: 29.2%; Neutral: 40.9%). Although they might have noticed that these interpretations were not correct, semantic representations of unrelated morphemes can be activated in their representations, without being suppressed in their text memory. In addition, given the low availability of distantly located linguistic clues (Pulido, 2009; Wesche & Paribakht, 2010) and the use of directive context in the experiment allow us to conclude that such misinterpretations would not be revised as the learners proceed through a text.

In sum, opaque words in a text would affect the text comprehension in two ways: (a) readers' distortion of contextual meanings in the process of inferencing, and (b) the lasting

semantic representation of morpheme-based meanings as part of text comprehension. These effects caused by opaque words can further impair a more global comprehension of the text (Bensoussan & Laufer, 1984; Laufer, 1989).

5.2 Use of Morphological and Contextual Information in Lexical Inferencing by Beginner-Level Learners

Study 2 included two experiments (Experiments 4, 5) to answer the following research questions regarding the beginner-level Japanese EFL learners' use of morphological and contextual information in lexical inferencing.

RQ 4-1: To what degree do beginner-level Japanese EFL learners understand the sematic relationship between morphological and contextual information?

RQ 4-2: Are beginner-level Japanese EFL learners able to narrow down the meanings of unknown words according to context informativeness?

RQ5-1: Does lexical inference training using an inference worksheet support the lexical inference performance of beginner-level Japanese EFL learners?

RQ5-2: Does lexical inference training improve the lexical inference skills of beginner-level Japanese EFL learners?

Based on the findings from previous studies, it was apparent that beginner-level learners frequently made incorrect word-based interpretations (e.g., Hamamda, 2014). Therefore, the experiments focused on identifying their particular difficulties in the appropriate use of the two sources of information according to the semantic transparency of morphemes, and examine the effects of inference training based on them. In order to determine the focus of inference training of Experiment 5, Experiment 4 examined whether beginner-level Japanese EFL learners could

understand the semantic in/consistency between morphemes and the context when they were explicitly asked to judge the relationship (RQ4-1) and their ability to use contextual information for inferencing (RQ4-2). The participants undertook the on- and off-line versions of semantic consistency judgment tasks. The results of accurate response rates showed that beginner-level learners could understand the relationship in terms of accuracy in both on- and off-line tasks. However, lower performance for transparent condition in the on-line task, and lower confidence ratings in the off-line judgments in all conditions, indicated the possibility that they might not have understood the context. The results of the reaction times for the on-line task showed that they could activate more semantics of unknown words from informative contexts than neutral ones, as suggested by the faster reaction times for informative conditions.

Experiment 5 investigated the effects of one-shot inference training on the performance of lexical inferences. Based on the findings of Experiment 4, the training mainly focused on the use of linguistic clues in inferring the meanings of unknown words. In addition, a step-by-step inference worksheet was prepared to reduce the cognitive complexity of performing multiple processes involved in it. The participants, further divided into three groups (i.e., Upper, Middle, Lower groups), took the pre-test, and then worked on the inference training. One week after the treatment, they took the post-test to see if the effects of the training were retained and if they applied their learning to normal lexical inferencing. The performance comparison between the pre-test and inference training showed the immediate effects of the training; the participants in all proficiency groups made fewer literal interpretations of opaque words. The Middle and Upper groups made more contextually appropriate interpretations, though the proportion of increase was rather small, while in the case of the Lower group, there was no increase. Furthermore, the above positive effects of the training were not observed in the post-test.

Given that the focus of Study 2 lies in supporting the lexical inferences of beginner-level EFL learners, the results of the two experiments could not produce positive results. Therefore,

the discussion focuses on the relationship between the two processes involved in appropriate interpretations of opaque words (i.e., perception of the semantic relationship between the two sources, generation of contextually appropriate meanings) and the performance of the present participants, in order to gain insights into how we should help them with future practice.

Understanding of semantic relationship between morphemes and the context

Previous studies have shown frequent incorrect word-based interpretations of opaque words among less skilled readers. They were attributed to insufficient monitoring of the two sources (e.g., Hamada, 2014) and their piece-by-piece reading style (Oakhill et al., 2016).

In this regard, both Experiments 4 and 5 had the participants explicitly consider the semantic relationship between the two sources through the semantic consistency judgment task, where they were asked to judge the semantic in/consistency between the morphemes and contexts. Table 5.2 presents the correct response rates of semantic consistency judgment tasks in informative contexts in Experiments 4 and 5.

Table 5.2

Correct Response Rates for the Semantic Consistency Judgment Task for Informative Context (%) and the Participants' Scores for the Reading Proficiency Test in Experiments 4 and 5

		Semantic consistency judgment task				Proficiency test	
		Transparent		Opaque			
Experiment	*n*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Ex. 4 (On-line)	21	76.96	16.79	83.65	12.57	9.71	2.12
Ex. 4 (off-line)	21	86.31	11.79	83.93	14.94	9.71	2.12
Ex.5 (Upper)	37	63.12	26.48	83.21	18.28	10.73	1.54
Ex. 5 (Middle)	41	65.64	21.92	62.46	29.47	6.93	0.72
Ex. 5 (Lower)	41	65.17	24.63	54.95	25.49	4.00	1.02

It was found that the participants in Experiment 4 (both the on- and off-line tasks) and

the Upper group in Experiment 5 had more correct responses than the Middle and Lower group participants in Experiment 5. From this finding, it can be argued that certain reading proficiency level is required for correctly understanding the relationship between morphological and contextual information even when the readers attend to both the sources.

However, smaller proportion of correct response rates for transparent words among the Upper group in Experiment 4 (62.5%) than in Experiment 4 in the off-line task (86.3%) seemed strange given the similar mean scores of the reading proficiency test (Experiment 4: 9.2; Upper in Experiment 5: 10.7). The possible reason for this is that the subsequent inferencing process might have changed the standard of judgment. In Experiment 4, the task of the participants was to only judge the semantic in/consistency between the two sources. In contrast, the participants in Experiment 5 had to infer the meaning of the target words after the semantic judgment; therefore, they might set a stricter standard for transparent words. This was because the "Yes" response (literal meanings the morphemes fit to the context semantically) in the inference training (Step 2) automatically led to the adoption of literal interpretations they came up with in the previous step (Step 1). Thus, it was possible that they were afraid of making incorrect judgments and subsequent inappropriate interpretations. This was supported by higher proportions of MBI for transparent words than that of appropriate response of consistency judgment, which was observed in all groups. The participants made literal interpretations in the subsequent steps (i.e., Steps, 3, 4) even after they had judged the meanings to be semantically impossible in Step 2.

On the other hand, the smaller proportions of correct responses of the Middle (Transparent: 65.6%; Opaque: 62.5%) and Lower groups (Transparent: 65.2%; Opaque: 55.0%) for both transparent and opaque words showed that they had difficulty in understanding the semantic relations. Given that they were explicitly instructed to consider the two sources and judge the relation, the difficulty for these learners can be related to their insufficient contextual

understanding. The lack of reading proficiency prevented them from constructing sufficient context-based interpretations. As a result, they had to base their judgment on such inaccurate representations, which resulted in the smaller proportion of correct responses.

Overall, the findings suggest that even beginner-level EFL learners, with a certain degree of reading proficiency, could at least understand the semantic relationship between morphemes and the context when their attention was directed to this relationship.

Generation of contextually appropriate meanings

Based on the result of Experiment 4, the inference training used in Experiment 5 had the participants explicitly consider the two sources of information and made them infer the meanings of target words based on the semantic relations (i.e., semantic transparency). In addition, the training used only informative contexts based on the result of Experiment 4 to enable them to narrow down to more concrete ideas about target word meanings from the context. However, the results of the inference training did not produce effects as positive as were expected, even when the participants were provided the inference sheet that described how to use the morphological and contextual information in the course of inferencing step by step.

Figure 5.2 shows the overall performance of the inference training in Experiment 5. It was found that the difficulties lie in both (a) perception of the semantic relations (especially for the Middle and Lower groups) and (b) generation of contextually appropriate meanings (all groups). Since they were explicitly asked to engage in the two processes by judging the semantic inconsistency (i.e., Step 2) and generating contextually appropriate meanings either with parts of morpheme-based meanings (i.e., Step 3), or with fully context-based meanings (i.e., Step 4) in the training, the proportions shown in the figure should, to some degree, reflect their performance throughout the processes. The reason why there are discrepancies between the failure in the former process (Upper: 16.8%; Middle: 37.5%; Lower: 45.1%) and the

proportions of MBI (Upper: 19.6%; Middle: 34.5%; Lower: 36.6%), was that (a) they sometimes correctly detected the inconsistencies but later adopted the literal interpretations in the subsequent steps, and (b) failed the consistency judgements despite their answers that were produced in Step 1 (literal interpretations of morphemes) not being morpheme-based interpretations.

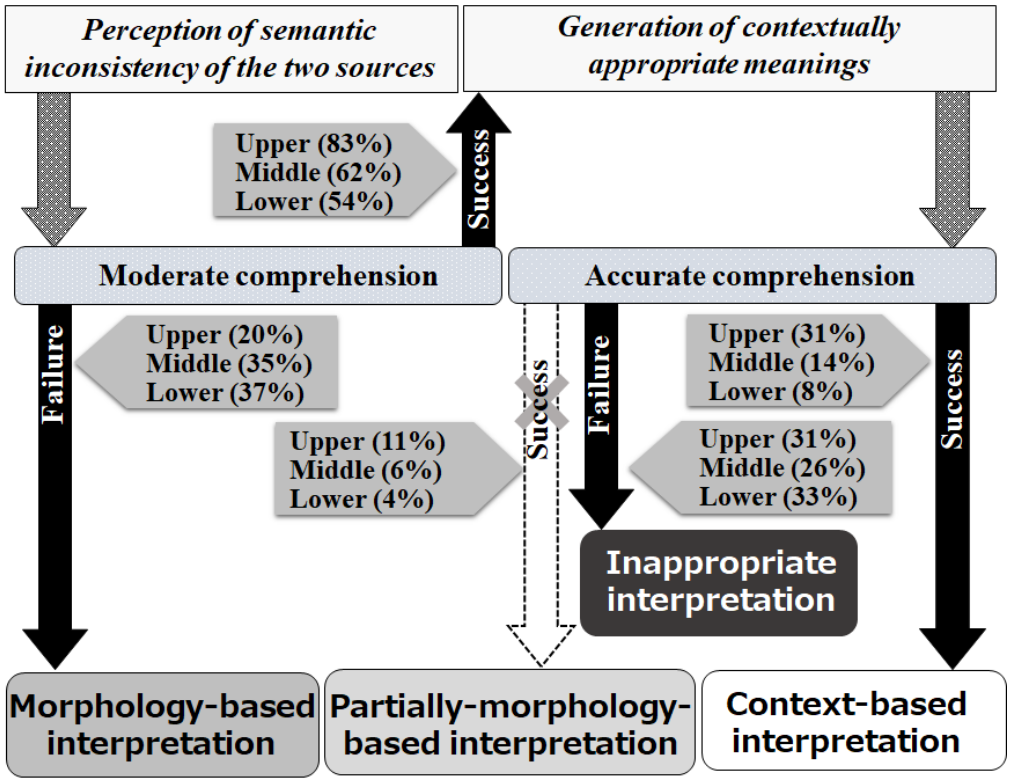


Figure 5.2. Relationship among the processes involved in inferencing opaque words, interpretation types, the level of required comprehension of the context, and the performance of the participants in Experiment 5 in the inference training. The gray x indicates that the process was hardly seen.

In all groups, the participants made inappropriate interpretations in many cases (Upper: 30.7%; Middle: 26.2%; Lower: 33.0%). Therefore, the correct understanding of the semantic

relation does not lead to learners succeeding at generating appropriate meanings. This discrepancy indicated the difference in context interpretations required for the two processes. While successful perception of the semantic relation can be achieved if the participants can understand superficial ideas of contexts and verify them with literal interpretations of target words (i.e., moderate comprehension), success in generation of appropriate meanings would require accurate interpretations since they need to narrow down the semantic context to fit the given context. Therefore, different participants among beginner-level learners experienced differing degrees of problems in the two processes. Moreover, the proportion of PMI was rather small in all groups, and it decreased in the Upper group. Since they were encouraged to consider integrated interpretations before forming fully context-based meanings, these results suggest that it is difficult for them to integrate semantics from the two sources.

However, this result slightly contradicts the result of Experiment 4 that reaction times of the semantic consistency judgment task words were shorter when the preceding context was an informative context rather than a neutral context. This means that they were able to activate more specific semantics about the upcoming target words, and the representations allowed them to make faster responses. Thus, if they are able to narrow down the meaning from the context, making contextual meanings is not so demanding. One possible reason for this was that the activated semantic representations were not concrete or specific enough to facilitate the generation of meanings. Since the informative contexts included more clue words, it was possible for the participants to activate related concepts without understanding the context accurately.

In addition, it was also possible that the presence of target words made it difficult for them to consider fully context-based meanings. In Experiment 4, the target words were presented after they had read the context sentence; thus, they were able to consider, and possibly predict, the upcoming words' meanings solely from contexts. On the other hand, in the training

of Experiment 5, target words and contexts were presented at once, and they were asked to write down the morpheme-based meaning on the worksheet (i.e., Step 1); as a result, the semantic information about the morphemes was activated in the minds of the participants. Therefore, the morpheme-based representation might have prevented them from fully considering context-based meanings.

This possibility of interference of morpheme-based meanings may explain why the beneficial effects of the inference training (fewer MBI, exploration of contextually appropriate meanings) observed in all groups of participants were not retained in the post-test. In the training, the participants were able to follow the inference procedure consisting of four steps one at a time the thanks to the worksheet. However, the absence of such aid in the post-test made it difficult for them to replicate the process in their minds, even when they remembered the content of the procedure; thus, it was cognitively demanding to consider the contextually appropriate meanings, simultaneously suppressing the activated morpheme-based meanings, especially for the learners with limited processing skills.

In conclusion, beginner-level learners' use of morphological and contextual information in lexical inferencing is limited, and as compared to intermediate-level learners, it can be affected to a greater degree by the level of text comprehension of each individual.

Chapter 6

Conclusion

6.1 Major Findings of this Study

This study aimed to examine how Japanese EFL learners with intermediate- or beginner-level proficiency process and interpret unknown words with semantically transparent and opaque morphological structures. The main findings of the study can be summarized as follows.

With regard to intermediate-level learners, although they are sensitive to the semantic relationship (in/consistency) between morphemes and context while reading, it is less likely for them to generate contextually appropriate meanings for opaque words. However, when the meanings of unknown words are strongly required for task completion (i.e., in lexical inferencing tasks), they strive to make interpretations that are semantically plausible in the context, flexibly using linguistic clues and monitoring their inferential processes. In doing so, the degree of contextual support may affect the types of information they use for interpretations (both morphemes and context, fully context-based). However, they sometimes revise or distort contextual meanings so that literal interpretations of morphemes of opaque words make sense within the context. In addition, incorrect morpheme-based interpretations are less likely to be revised as they proceed through a text. Overall, intermediate-level learners are equipped with the necessary linguistic and metacognitive skills to resolve the conflicts regarding morphemes and context; however, motivational and textual factors could affect how deeply they process unknown words and resulting interpretations.

In contrast, beginner-level learners' effective use of the two sources of information is limited. They face difficulties both in perceiving the semantic relationship between morphemes and context and the subsequent meaning generation process. Although inference training employing a worksheet that describes how to use linguistic clues may have helped them explore

contextually appropriate meanings, discarding inappropriate literal interpretations of morphemes, it led to only a small increase in appropriate interpretations by learners due to their insufficient comprehension skills. In addition, they seemed to face difficulty in operating multiple inferential processes without the worksheet one week after the training possibly due to their limited processing skills.

6.2 Limitations and Suggestions for Future Research

Although this study provides new insights into the use of morphological and contextual information when learners encounter unknown words and its relationship with learner proficiency, the present findings have some limitations and unresolved issues that should be addressed in future research. We will now list these points, proposing alternative solutions and discussing their potential effects.

First, this study consistently measured participants' interpretations of unknown words in an open-ended format, since we were particularly interested in how learners generate meanings by themselves after noticing semantic inconsistencies. Therefore, we asked participants to translate the context (Experiment 1), provide answers to the comprehension questions (Experiment 2), and infer the meanings of words (Experiments 3 and 5). We then analyzed the information sources they used for interpretations from their responses (e.g., morphemes, context). However, there may be a gap between the actual semantic representations of words in the readers' minds and their responses, which may not reflect their relatively unconscious and unverbalizable representations. For example, participants often inappropriately interpreted opaque words literally even though they perceived that the interpretation was incorrect. Mixing these interpretations with typical overreliance on morphological information might lead to underestimating the interpretation process. It is possible to avoid this situation by combining a multiple-choice format (e.g., morpheme-based vs. context-based options) with a response time

measure (Brusnighan & Folk, 2012, Experiment 2). Tasks requiring faster response would rule out the effects of a strategic process and could tap into activated semantic representations of the words even ones at slight and unconscious level.

Second, this study employed short contexts consisting of either one (Experiments 1, 3, 4, and 5) or two sentences (Experiment 2); as a result, participants could easily recognize the entire context at once. However, in longer texts, readers are required to search for contextual information in both preceding and succeeding sentences. This makes noticing inconsistencies more difficult, although this problem was not quite pervasive in this study. In addition, it is implied that the context length affects readers' attention to words: readers tend to attend to unknown words more in short contexts than in longer texts (Elgort & Warren, 2014). Thus, the presence of target words may have been more salient to the participants of this study, and they may have processed such words more deeply than in their normal reading. As a result, if longer texts were used, participants may have mistakenly identified the unknown words as known words, as has been observed in previous studies (Bensoussan & Laufer, 1984; Huckin & Bloch, 1993; Laufer, 1989).

The contexts used in this study were quite easy, especially for intermediate-level learners. This manipulation was done because one of our interests was to compare the performance of different levels of learners; thus, we adjusted the material levels to suit beginner-level learners. However, Ushiro et al. (2010), who employed moderately difficult sentences, found proficiency-related differences in learners with vocabularies of over 4,000 words. In their study, lower-proficiency groups were less able to flexibly change their interpretations than higher-proficiency groups. Therefore, if more complex material in terms of lexical levels and syntactic complexity had been adopted, even intermediate-level learners would have faced more difficulty in the two processes, especially in the meaning generation process. Therefore, we should keep in mind that proficiency-related differences observed in this study might not

necessarily be definitive. Rather, the interaction between materials and learner variables should probably be considered.

With regard to target words, this study used either transparent or opaque pseudo-compound words, the transparency of which was manipulated by the semantic in/consistency between literal interpretations of the words and their surrounding context (Brusnighan & Folk, 2012; Hamada, 2014). Hence, we did not entirely control the degree of transparency/opacity of the targets. However, the concept of semantic transparency by nature is not dichotomous; it is a continuum ranging from fully opaque to fully transparent (Libben & Jarema, 2006; Schäfer, 2018). Therefore, it was possible to perceive that the inconsistency of some items was relatively easy, while for other items it was relatively difficult. The findings of the norming studies of Experiments 1 and 2 (rated by graduate students majoring in English language education), and the rather small number of participants who did not monitor the two sources of information (Lack of monitoring; Experiment 3), which was observed only four times, indicated that proficient learners were sensitive to the slight difference of which a specific word belongs to which side of the continuum (transparent vs. opaque) in relation to the semantics of the two sources. On the other hand, for beginner-level participants, who showed less accurate responses for semantic consistency judgments, such variability in transparency would have affected their consistency judgments. Therefore, in addition to controlling these effects carefully, it is also interesting to examine how the degree of semantic transparency/opacity affects their perception of semantic in/consistency between morphemes and context.

Furthermore, it is necessary to address other lexical items with differing sematic relationships between literal and actual meanings. This is because the perception of semantic in/consistency and subsequent meaning generation can be affected by the semantic relationship between literal, word-based meanings and actual meanings. In this regard, participants found it relatively easy to recognize the present transparent or opaque words in the former process due

to the lack of semantic overlap between the two senses, while they experienced difficulty in the latter process because they had to suppress unnecessary morpheme-based meanings. Therefore, investigating both the abovementioned processes for lexical items with semantic relatedness, such as primary and secondary meanings of polysemous words and literal and figurative interpretations of idiomatic expressions, would provide a more comprehensive understanding of how we process and interpret unknown lexical items with discrepancies in literal and actual meanings in general.

Unlike intermediate-level learners, beginner-level learners showed limited effectiveness in using morphological and contextual information in inferencing, and the inference training in Experiment 5 did not provide adequate help to them. Therefore, more evidence is necessary to identify the cause of difficulty and explore better teaching practices. In this regard, the biggest limitation was the lack of longitudinal intervention: the one-shot inference training of Experiment 5 was deemed insufficient to acquire the inferential process. Therefore, we should conduct repeated practices and track how they can help develop inferential skills over time.

In addition, beginner-level learners' poor inference performance was attributed to their inaccurate understanding. However, we did not directly measure their context interpretations. In this regard, it was discussed that initial perception of the semantic relationships and subsequent generations of contextually appropriate interpretations require moderate and accurate comprehension of context, respectively. Therefore, in order to verify and extend the findings, examining the relationship between context interpretation and inference outcomes will be useful to estimate the exact degree of context understanding required for both the processes, and to identify the level of comprehension in which these learners face difficulties, such as lexical or propositional level.

Finally, in terms of vocabulary learning from context, due to the semantic conflicts between part (morphemes) and whole word meanings, incorporating the lexical knowledge of

opaque words from contextual exposure into one's lexicon can be more demanding, compared to monomorphemic and semantically transparent words. Thus, it is worthwhile to examine how lexical knowledge of these words develops from increased exposure (e.g., Webb, 2007). Also, examining the effects of initial intentional learning of these items and how it influences subsequent lexical access in context can provide us with insight on how teachers should deal with such problematic words. This is particularly instrumental in EFL settings, where a combination of intentional and incidental learning plays a key role in vocabulary learning (Kadota & Ikemura, 2006).

6.3 Pedagogical Implications

This study identified the exact source of difficulties and conditions involved in the successful understanding of unknown words and its relationship with learner proficiency. These findings have pedagogical implications regarding the treatment of unknown words, especially semantically opaque words. Given the clear proficiency-related differences found in this study, we will provide implications that are necessary for intermediate- or more advanced-level and beginner-level learners separately.

Implications for intermediate- or more advanced-level learners

The first implication is related to the improvement of intermediate- or advanced-level learners' inferential skills. In this regard, participants generally made contextually appropriate interpretations for opaque words (Experiment 3). In addition, apart from a few exceptions in which they failed to notice semantic inconsistencies (i.e., Negligence of context, Lack of monitoring), other literal interpretations (i.e., Persistent reliance on morphemes, Modification of context, Elaboration of context) would be the result of not knowing how to resolve semantic conflicts. In Experiment 3, massed exposure to opaque words (16 times) may have induced

more conscious efforts than normal lexical inferencing; accordingly, these interpretations would appear more often in the experiment. Therefore, it is useful to inform learners that not all morphological information is reliable and that they should prioritize context-based meanings. Such explanations will solicit their conscious effort to reveal the meanings of words, focusing on the interpretations of unknown words rather than distorting contextual meanings.

When generating the meanings of words, participants tended to make fully context-based interpretations when a word was embedded in informative contexts, while interpretations that included both morphemes and contextual information were produced more frequently for neutral contexts (Experiments 1 and 2). These results reflect participants' sensitivity to contextual information. However, as in the case of polysemous words, idiomatic phrases, and semi-transparent words, there are situations in which readers could achieve more accurate interpretations by extending or using the meanings that are available from the lexical items. Given these linguistic facts, the inclusion of partial or associated meanings for interpretations should ideally be determined by whether semantic elements are possible within the context, not by the context quality. Therefore, teachers should provide opportunities to students to consider the information from both word and context meanings. Although combining morphological and contextual information, like semi-transparent words, is a cognitively demanding task (Hamada, 2014; Mori & Nagy, 1999), the abovementioned result of PMI showed that participants were able to accomplish it when asked to do so. In addition to semi-transparent words, only one of the two morphemes that is directly related to the actual meaning of the entire word, polysemous words, or idiomatic phrases, as well as semi-transparent words, are suitable materials for training because these items have metaphorical and figurative meanings for literal interpretations (Oakhill et al., 2016; Verspoor & Lowie, 2003). Inferencing training with these lexical items will help learners understand various types of relationships and contribute to raising their metalinguistic awareness.

In reading instruction, however, it is not always practical or effective to necessitate students to infer all unknown words. In addition, since lexical inferencing takes considerable time (Huckin & Coady, 1999), it is often recommended that readers skip or throw away unknown words if their meanings are not crucial for comprehension (Nation, 2013; Nuttal, 2003). However, since deceptively transparent words, including opaque words, are frequently perceived as known words (Laufer, 1989) and their word-based meanings are unrelated to their actual meanings, they may cause more substantial misinterpretations as compared with mere unknown words. Furthermore, as found in this study (Experiments 1 and 3), readers sometimes change the interpretations of context meaning rather than target words, suggesting that opaque words should be treated with attention to these potential effects on the comprehension of contexts surrounding target words. Therefore, if such problematic words are essential for text comprehension, asking questions regarding the meanings of unknown words in advance would motivate students to attentively process and infer the meanings of words. Even if these words belong to unimportant segments of a text, given the effects of other parts of the text, teachers should confirm the comprehension of words and surrounding contexts.

Implications for beginner-level learners

Unlike intermediate-level learners, beginner-level learners faced more difficulties inferring unknown opaque words; therefore, they need more help when they come across such words. Here, I propose methods for semantically opaque or deceptively transparent words in both short- and long-term perspectives, considering variation even among beginner-level learners.

Fist, learners who struggle to understand sentence meaning, as shown in the Lower group of Experiment 5, face difficulty in understanding the semantic relationship between the two sources of information. Therefore, it may be necessary for teachers to introduce problematic

words before reading to these learners. In that case, it is useful to examine whether new words have their familiar morphemes or formal similarity with their known words.

Second, when students are able to understand most of the ideas in a text, encouraging them to make inferences can be effective if the necessary help is provided. This is shown in the Upper and Middle groups of Experiment 5, in which participants were able to make more appropriate interpretations in the inference training. However, the large proportions of MBI in the pre-test of Experiment 5 showed that their text processing was more word-by-word and that they had failed to notice the semantic relationships. Therefore, necessitating them to explicitly attend to both the morpheme and context, as in the semantic consistency judgment task, will help them consider context meanings and notice that word-based information is incorrect. However, the large proportions of INI indicate that their context understanding would be insufficient to generate meanings by themselves. Therefore, help might be provided to such students by alleviating the burden of generating meanings, such as by providing them with multiple options. Although Hamada (2014) showed that the beginner group often made literal interpretations even in multiple-choice inference tests, once they were aware that morphological information was incorrect, they would explore other options, as seen in the training of Experiment 5. If teachers want students to generate meanings by themselves, asking them to infer fully context-based meanings would be a suitable first step since participants had difficulty in integrating the sources of information, as can be seen in the smaller proportion of PMI.

At the same time, teachers should develop students' inference skills so that they can infer the meanings of unknown words by themselves. In this regard, one of the most plausible reasons for the lack of success in the present inference-training was that participants did not have the necessary skills that would enable them to follow the content of training. Therefore, teachers need to improve students' basic linguistic skills that support accurate understanding through

reading instructions, such as accurate and efficient lexical access and syntactic parsing. These skills allow them to accurately understand the context and make successful context-based inferences, and also leave more cognitive resources for engaging in multiple processes involved in appropriate interpretations for opaque words.

In addition, the results of Experiment 5, which showed that the effects of the inference training were not retained when participants inferred without the worksheet one week after the training, suggest that more training is necessary to familiarize students with the inferential procedure. Therefore, by providing opportunities to students to practice the inferential procedure using a procedural aid, such as the worksheet used in this study, and gradually reducing support based on their performance, students will be able to infer unknown words considering morphemes and context by themselves.

6.4 Concluding Remarks

Success in lexical inferencing and reading comprehension are closely related in that lexical inferencing contributes to reading comprehension by filling in semantic gaps in a text, and good text comprehension provides context-based semantic foundations for lexical inferencing. However, the low success rates of L2 or EFL lexical inferencing, especially with low proficiency learners, might lead some researchers to believe that it is not an effective or efficient way to deal with unknown words as compared with other options, such as using a dictionary. However, readers' misinterpretations of unknown words, especially for opaque words, makes it difficult to use referential sources, which could also affect reading comprehension in general (Bensoussan & Laufer, 1984). Therefore, identifying the possible cause of such incorrect inferences and improving students' lexical inferencing skills is still important. Therefore, this study examined the factors that make interpretation difficult and explored possible interventions by focusing on the processes of inferencing as well as its

outcomes. Despite some limitations, the current research is significant because it identified the exact source of difficulties and conditions involved in effectively using morphological and contextual information in inferencing according to the proficiency levels of Japanese EFL learners.

References

Albrecht, J. E., & O'Brien, E. J. (1993). Updating a mental mode: Maintaining both local and global coherence. Journal of Experimental *Psychology: Learning, Memory, and Cognition, 19,* 1061–1070. doi:10.1037/0278-7393.19.5.1061

Bauer, L., Lieber, R., & Plag, I. (2013). *The Oxford reference guide to English morphology*. Oxford University Press.

Beck, I. L., McKeown, M. G., & McCaslin, E. S. (1983). Vocabulary development: All contexts are not created equal. *The Elementary School Journal*, *83,* 177–181. Retrieved from https://www.jstor.org/stable/1001501?seq=1#metadata_info_tab_contents

Bengeleil, N. F., & Paribakht, T. S. (2004). L2 reading proficiency and lexical inferencing by university EFL learners. *Canadian Modern Language Review*, *61*, 225–250. doi:10.3138/cmlr.61.2.225

Bensoussan, M., & Laufer, B. (1984). Lexical guessing in context in EFL reading comprehension. *Journal of Research in Reading*, *7*, 15–32. doi:10.1111/j.1467-9817.1984.tb00252.x

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. doi:10.1080/01638530701792826

Brusnighan, S. M. (2010). Semantic transparency and contextual strength in incidental vocabulary acquisition of novel compounds during silent reading: Evidence from eye movements and recall. Master's thesis, Kent State University. Retrieved from https://etd.ohiolink.edu/rws_etd/document/get/kent1286756820/inline

Brusnighan, S. M., & Folk, J. R. (2012). Combining contextual and morphemic cues is beneficial during incidental vocabulary acquisition: Semantic transparency in novel

compound word processing. *Reading Research Quarterly*, *47*, 172–190. doi:10.1002/RRQ.015

Cain, K., Oakhill, J., & Lemmon, K. (2005). The relation between children's reading comprehension level and their comprehension of idioms. *Journal of Experimental Child Psychology*, *90*, 65–87. doi:10.1016/j.jecp.2004.09.003

Cain, K., Towse, A. S., & Knight, R. S. (2009). The development of idiom comprehension: An investigation of semantic and contextual processing skills. *Journal of Experimental Child Psychology*, 102, 280–298. doi:10.1016/j.jecp.2008.08.001

Chern, C. L. (1993). Chinese students' word-solving strategies in reading in English. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 67–85). Norwood, NJ: Ablex.

Clarke, D., Nation, I.S.P. (1980). Guessing the meaning of words from context: Strategy and techniques. *System*, *8*, 211–220. doi:10.1016/0346-251X(80)90003-2

Coady, J. (1993). Research on ESL/EFL vocabulary acquisition: Putting it in context. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 3–23). Norwood, NJ: Ablex.

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671–684. doi:10.1016/S0022-5371(72)80001-X

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. doi:10.1017/S0272263197003021

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language*, *27*, 429–446. doi:10.1016/0749-596X(88)90066-6

Elgort, I. (2017). Incorrect inferences and contextual word learning in English as a second language. *Journal of the European Second Language Association*, *1*, 1-11. doi:10.22599/jesla.3

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. doi:10.1111/lang.12052

Elston-Güttler, K. E., & Friederici, A. D. (2005). Native and L2 processing of homonyms in sentential context. *Journal of Memory and Language*, *52*, 256–283. doi:10.1016/j.jml.2004.11.002

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

Frantzen, D. (2003). Factors affecting how second language Spanish students derive meaning from context. *The Modern Language Journal*, *87*, 168–99. doi:10.1111/1540-4781.00185

Fraser, C. A. (1999). Lexical processing strategy use and vocabulary learning through reading. *Studies in second language acquisition*, *21*, 225–241. doi:10.1017/S0272263199002041

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

Haastrup. K. (1991). *Lexical inferencing procedures or talking about words: Receptive procedures in foreign language learning with special reference to English.* Tübingen, Germany: Gunter Narr.

Hamada, A. (2013). Lexical inference specificity and its activation level: Effects of contextual constraint and L2 reading proficiency. *ARELE* (*Annual Review of English Language Education in Japan)*, *24*, 93–108. doi:10.20581/arele.24.0_93

Hamada, M. (2014). The role of morphological and contextual information in L2 lexical inference. *The Modern Language Journal*, *98*, 992–1005. doi:10.1111/modl.12151

Haynes, M. (1993). Patterns and perils of guessing in second language reading. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 46–66). Norwood, NJ: Ablex.

Horiba, Y. (2012). Word knowledge and its relation to text comprehension: A comparative study of Chinese-and Korean-speaking L2 learners and L1 speakers of Japanese. *The Modern Language Journal*, *96*, 108–121. doi:10.1111/j.1540-4781.2012.01280.x

Horiba, Y. (2013). Task-induced strategic processing in L2 text comprehension. *Reading in a Foreign Language*, *25*, 98–125. Retrieved from https://files.eric.ed.gov/fulltext/EJ1015753.pdf

Hu, H. M., & Nassaji, H. (2012). Ease of inferencing, learner inferential strategies, and their relationship with the retention of word meanings inferred from context. *Canadian Modern Language Review*, *68*, 54–77. doi:10.3138/cmlr.68.1.054

Hu, H. M., & Nassaji, H. (2014). Lexical inferencing strategies: The case of successful versus less successful inferencers. *System*, *45*, 27–38. doi:10.1016/j.system.2014.04.004

Hu, M. H., & Nation, I. S. P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, *13*, 403–430. Retrieved from http://nflrc.hawaii.edu/rfl/PastIssues/rfl131hsuehchao.pdf

Huckin, T., & Bloch, J. (1993). Strategies for inferring word-meanings in context: A cognitive model. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 153–178). Norwood, NJ: Ablex.

Huckin, T., & Coady, J. (1999). Incidental vocabulary acquisition in a second language. *Studies in Second Language Acquisition*, *21*, 181–93.

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

Hulstijn, J. H. (1993). When do foreign-language readers look up the meaning of unfamiliar words? The influence of task and learner variables. *The Modern Language Journal*, *77*, 139–147. doi:10.2307/328937

Hulstijn, J. H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning*, *51*, 539–558. doi:10.1111/0023-8333.00164

Hyönä, J. (2015). Are polymorphemic words processed differently from other words during reading? In A. Pollatsek., & R. Treiman (Eds.), *The Oxford handbook of reading* (pp. 114–128). Oxford University Press.

Ishikawa, S., Uemura, T., Kaneda, M., Shimizu, S., Sugimori, N., Tono, Y., . . . Murata, M. (2003). *Japan Association of College English Teachers 8000: JACET list of 8000 basic words.* Tokyo, Japan: JACET.

Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics*, *25*, 603–634. doi:10.1017/S0142716404001298

Jiang, N. (2007). Selective integration of linguistic knowledge in adult second language learning. *Language Learning*, *57*, 1–33. doi:10.1111/j.1467-9922.2007.00397.x

Jiang, N. (2012). Conducting reaction time research in second language studies. New York, NY: Routledge.

Kadota, S., & Ikemura, D. (Eds). (2006). *Eigo goi shidou handbook* [Handbook of English vocabulary instruction]. Tokyo, Japan: Taishukan Publishing.

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist*, *49*, 294–303.

Koda, K. (2000). Cross-linguistic variations in L2 morphological awareness. *Applied Psycholinguistics*, *21*, 297–320. doi:10.1017/S0142716400003015

Laufer, B. (1989). A factor of difficulty in vocabulary learning: Deceptive transparency. *AILA Review*, *6*, 10–20.

Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical Threshold Revisited: Lexical Text Coverage, Learners' Vocabulary Size and Reading Comprehension. *Reading in a Foreign Langu*age, *22*, 15–30. Retrieved from http://nflrc.hawaii.edu/rfl/April2010/articles/laufer.pdf

Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics, 22*, 1–26. doi:10.1093/applin/22.1.1

Levelt, W. J. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levorato, M. C., & Cacciari, C. (1995). The effects of different tasks on the comprehension and production of idioms in children. *Journal of Experimental Child Psychology*, *60*, 261–283. doi:10.1006/jecp.1995.1041

Levorato, M. C., & Cacciari, C. (1999). Idiom comprehension in children: Are the effects of semantic analysability and context separable? *European Journal of Cognitive Psychology*, *11*, 51–66. doi:10.1080/713752299

Libben, G., & Jarema, G. (Eds.). (2006). *The representation and processing of compound words*. Oxford University Press.

Mochizuki, M., & Aizawa, K. (2000). An affix acquisition order for EFL learners: An exploratory study. *System*, *28*, 291–304. doi:10.1016/S0346-251X(00)00013-0

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. doi:10.1093/applin/12.3.249

Mori. Y. (2002). Individual differences in the integration of information from context and word parts in interpreting unknown kanji words. *Applied Psycholinguistics*, *23*, 375–397.

doi:10.1017/S0142716402003041

Mori, Y., & Nagy, W. (1999). Integration of information from context and word elements in interpreting novel kanji compounds. *Reading Research Quarterly*, *34*, 80–101. doi:10.1598/RRQ.34.1.5

Morishima, Y. (2013). Allocation of limited cognitive resources during text comprehension in a second language. *Discourse Processes*, *50*, 577–597. doi:10.1080/0163853X.2013.846964

Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly*, *19*, 304–330. doi:10.2307/747823

Nakagawa, C. (2006). The effects of morphological and contextual clues on EFL readers' lexical inference. *ARELE* (*Annual Review of English Language Education in Japan)*, *17*, 151–160. doi:10.20581/arele.17.0_151

Nassaji, H. (2003). L2 vocabulary learning from context: Strategies, knowledge sources, and their relationship with success in L2 lexical inference. *TESOL Quarterly*, *37*, 645–670. doi:10.2307/3588216

Nassaji, H. (2006). The relationship between depth of vocabulary knowledge and L2 learners' lexical inferencing strategy use and success. *The Modern Language Journal*, *90*, 387–401. doi:10.3138/cmlr.61.1.107

Nation, I. S. P. (2009). *Teaching vocabulary: Strategies and techniques.* Cambridge University Press.

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

Nuttall, C. (2005). *Teaching reading skills in a foreign language*. Oxford, UK: Macmillan Education.

Oakhill, J., Cain, K., & Nesi, B. (2016). Understanding of idiomatic expressions in context in

skilled and less skilled comprehenders: Online processing and interpretation. *Scientific Studies of Reading*, *20*, 124–139. doi:10.1080/10888438.2015.1092973

Obunsha. (2000). *Eiken jun ni kyu zen mondai-shu* [The collection of previous questions for the pre-second grade Eiken test]. Tokyo: Author.

Obunsha. (2001a). *Eiken ni kyu zen mondai-shu* [The collection of previous questions for the second grade Eiken test]. Tokyo: Author.

Obunsha. (2001b). *Eiken san kyu zen mondai-shu* [The collection of previous questions for the third grade Eiken test]. Tokyo: Author.

Obunsha. (2008). *Eiken jun ikkyu zen mondai-shu* [The collection of previous questions for the pre-first grade Eiken test]. Tokyo: Author.

Paribakht, T. S., & Wesche, M. (1999). Reading and "incidental" L2 vocabulary acquisition. *Studies in Second Language Acquisition*, *21*, 195–224.

Pollatsek, A., Bertram, R., & Hyönä, J. (2011). Processing novel and lexicalised Finnish compound words. *Journal of Cognitive Psychology, 23,* 795–810. doi:10.1080/20445911.2011.570257

Pollatsek, A., Slattery, T. J., & Juhasz, B. (2008). The processing of novel and lexicalised prefixed words in reading. *Language and Cognitive Processes*, *23*, 1133–1158. doi:10.1080/01690960801945484

Pulido, D. (2003). Modeling the role of second language proficiency and topic familiarity in second language incidental vocabulary acquisition through reading. *Language Learning*, *53*, 233–284. doi:10.1111/1467-9922.00217

Pulido, D. (2009). How involved are American L2 learners of Spanish in lexical input processing tasks during reading? *Studies in Second Language Acquisition*, *31*, 31–58. doi:10.1017/S0272263109090020

Schäfer, M (2018). *The semantic transparency of English compound nouns*. Berlin, Germany:

Language Science Press.

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Palgrave Macmillan: London, UK.

Shintomi, H., Yamane, K., Watanabe, S. (1998). *Gyaku-biki eigo meishi fukugougo jiten* [An inverted dictionary of English nominal compounds]. Tokyo, Japan: Hokuseido Press.

Swanborn, M. S., & de Glopper, K. (1999). Incidental word learning while reading: A meta-analysis. *Review of Educational Research*, *69*, 261–285. doi:10.3102/00346543069003261

Ushiro, Y., Hamada, A., Mori, Y., Hosoda, M., Tada, G., Kamimura, K., & Okawara, N. (2018). Goal-oriented L2 reading processes in maintaining the coherence of narrative comprehension. *JACET Journal*, *62*, 109–128.

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*, 1–19.

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. *ARELE (Annual Review of English Language Education in Japan)*, *21*, 161–170. doi:10.20581/arele.21.0_161

Ushiro, Y., Mori, Y., Hosoda, M., Tanaka, N., Dowse, E., Tada, G., & Nakagawa, H. (2016). How do Japanese EFL readers maintain coherence in narrative memory? *ARELE (Annual Review of English Language Education in Japan)*, *27*, 81–96. doi:10.20581/arele.27.0_81

Ushiro, Y., Nahatame, S., Hasegawa, Y., Kimura, Y., Hamada, A., Tanaka, N., Hosoda, M., Mori, Y. (2016). Maintaining the coherence of situation models in EFL reading: Evidence from eye movements. *JACET Journal*, *60*, 37–55.

van den Broek, G. S., Takashima, A., Segers, E., & Verhoeven, L. (2018). Contextual richness and word learning: Context enhances comprehension but retrieval enhances retention. *Language Learning*, *68*, 546–585. doi:10.1111/lang.12285

Verspoor, M., & Lowie, W. (2003). Making sense of polysemous words. *Language Learning*, *53*, 547–586. doi:10.1111/1467-9922.00234

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*, 130–183. Retrieved from http://nflrc.hawaii.edu/rfl/October2003/waring/waring.html

Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, *28*, 46–65. doi:10.1093/applin/aml048

Webb, S. (2008). The effects of context on incidental vocabulary learning. *Reading in a Foreign Language*, *20*, 232–245. Retrieved from https://pdfs.semanticscholar.org/10b4/7af9fafc38b638ce1c18cb2d4e4890ac9e19.pdfF

Webb, S., & Nation, I.S.P. (2017). *How vocabulary is learned*. Oxford University Press.

Wei, Z. (2015). Does teaching mnemonics for vocabulary learning make a difference? Putting the word method and the word part technique to the test. *Language Teaching Research*, *19*, 43–69. doi:10.1177/1362168814541734

Wesche, M. B., & Paribakht, T. S. (2010). *Lexical inferencing in a first and second language: cross-linguistic dimensions.* Bristol, England: Multilingual Matters.

Appendices

Appendix 1

A list of Target Words, and Their Creation Procedure

Original words	Target words	Original words	Target words
dry nurse	babyhelper	locker room	lockerbox
battercake	breadcake	lunchtime	lunchhour
work bag	businessbag	meatball	meatblock
campsite	campspot	gift certificate	moneyticket
NA	designbook	morning bird	morninglife
NA	floorcover	night work	nightjob
airplane	flymachine	paint brush	paintstick
NA	fruitsweet	price war	pricebattle
earthquake	groundshake	safety zone	safetyarea
watchman	guardperson	shoreline	sealine
stray bullet	gunball	footmark	shoesmark
pocket	handknife	sunrise	sunclimb
hometown	homecity	railroad	trainroad
desk clerk	hotelstaff	freight station	trainstop
house coat	housewear	footpath	walkroad
telephone pole	lightpole	swimwear	waterwear

Note. NA implies that the target word was adapted from Brusnighan and Folk (2010).

Appendix 2

Definitions and Examples of Interpretation Categories and Their Relationships With Appropriateness, According to Semantic Transparency

Category	Definition and examples	Transparent	Opaque
MBI	Definition: Literal translations of both two morphemes Example: Jerry went to a big shopping center to buy a **waterwear** three days ago.→ 水着 (a swimsuit)	○	×
PMI	Definition: Translations where part of the morphemes' semantic information was deleted or modified, some information was added, or the interpretation involved their associated meanings. The meanings are semantically appropriate in the contexts. Example: Dan was looking for the nearest **fruitsweet** from his house. → 果物屋 (a fruit store); For her birthday present, Emi wanted a **trainroad** she had seen on TV. → 線路の模型 (a model of a railroad); After arriving at the office, he realized he had left his pricebattle on the train. → 財布 (a wallet)	○	○
CBI	Definition: Translations whose meanings were appropriate in the contexts, and no morphemic information was included. Example: To prepare for the art class, the student bought a **groundshake** at the store. → パレット (a pallet)	○	○
INI	Definition: Translations whose meaning was inappropriate in the contexts. It covers cases where part of the morpheme information was included. Example: After arriving at the office, he realized he had left his **pricebattle** on the train. → 時間 (time); The workers had to build a **meatblock** within one week. → 肉 (meat)	×	×
None	Definition: No translation corresponding to target words was included. (*I don't know* response in Experiment 2)	×	×

Appendix 3

Materials Used in Experiments 1, 3, 4, and 5

Practice sentences (Experiment 1)

(1) Wendy paid for the bag, but left it on the counter.

CQ. Did she forget to take the bag with her? (Yes)

(2) Roger was the owner of the local newspaper company at that time.

(3) The picture on the postcard was of a church in France.

CQ. Was the picture taken in Mexico? (No)

Practice sentences (Experiment 3)

(1) The students had three tests today, so they wanted to **stopket** tonight.

(2) He was not ill, and of course the beds in the **ancon** are for ill people.

(3) They tried to make their **sadmesk** friend feel better.

Practice sentences (Experiment 4)

(1) The student forgot to bring his **schooltext**. (YES)

(2) Ronald wanted to do **headpain**. (NO)

(3) Betty's mother asked her to buy **milkdrink**. (YES)

Experimental sentences (Experiments 1, 3, 4, 5)

(1) groundshake / designbook

Inf:	Sara lost her home because of the big **groundshake** / **designbook** five years ago.
Neu:	Shorn didn't feel the **groundshake** / **designbook** on that night.
Inf:	To prepare for the art class, the student bought a **designbook** / **groundshake** at the store.
Neu:	The girl was happy after she bought a **designbook** / **groundshake** at the shop.

Note. The first and second target words (boldfaced) served as transparent and opaque target words, respectively. Inf = informative context and Neu = neutral context.

(2) lunchhour / waterwear

Inf:	Kevin bought some food to eat during his **lunchhour** /**waterwear** near the office.
Neu:	Ellis decided to hold a meeting during the **lunchhour** /**waterwear** to discuss the problem.
Inf:	During the summer sale, the sport shop sold a **waterwear** / **lunchhour** at a low price.
Neu:	Jerry went to a big shopping center to buy a **waterwear** / **lunchhour** three days ago.

(3) trainroad /paintstick

Inf:	The station worker found something strange on the **trainroad** / **paintstick** in the morning.
Neu:	The workers planned to build a **trainroad** / **paintstick** in the city.
Inf:	Ken went to an art shop and bought a **paintstick** / **trainroad** that he wanted.
Neu:	For her birthday present, Emi wanted a **paintstick** / **trainroad** she had seen on TV.

(4) gunball / hotelstaff

Inf:	The police checked the dead man's body and found a small **gunball** / **hotellstaff** near his heart.
Neu:	Ronald had never bought a **gunball hotellstaff** in his life.
Inf:	Tim got his room key and gave his bags to the **hotelstaff** / **gunball** at the counter.
Neu:	Sarah talked with a **hotelstaff** / **gunball** on the phone.

(5) lightpole / meatblock

Inf:	Tim was able to find his lost phone in the dark thanks to the **lightpole** / **meatblock** along the street.
Neu:	The workers had to build a **lightpole** / **meatblock** within one week.
Inf:	To cook steak for her family, Jessica bought a **meatblock** / **lightpole** at the supermarket.
Neu:	Demi forgot to buy a **meatblock** / **lightpole** that her friend had asked.

(6) guardperson / handknife

Inf:	To protect the important guest, the city decided to hire a **guardperson** / **handknife** at the meeting.
Neu:	Ellen thought that the woman was a **guardperson** / **handknife** from her clothes.
Inf:	Greg was cooking outside and he used a **handknife** / **guardperson** on that day.
Neu:	Susan didn't know how to use the **handknife** / **guardperson** when she was 12.

(7) campspot / morninglife

Inf:	The young group enjoyed a barbecue at a **campspot** / **morninglife** for more than three hours.
Neu:	Tom and Cathy decided to go to the **campspot** / **morninglife** during their vacation.
Inf:	To keep healthy, Steve had been doing **morninglife** / **campspot** since last Sunday.
Neu:	Cindy tried to do **morninglife** / **campspot** but soon gave up.

(8) moneyticket / trainstop

Inf:	The man was in trouble when the shop didn't accept his **moneyticket** / **trainstop** at the counter.
Neu:	Jill's friend gave him a **moneyticket** / **trainstop** as a birthday present.
Inf:	For his work, Matt moved into a house near the **trainstop** / **moneyticket** with his wife.
Neu:	Rolla told a foreigner the way to the **trainstop** / **moneyticket** after finishing work.

(9) fruitssweet / safetyarea

Inf:	For dessert, Jane ate a delicious **fruitssweet** / **safetyarea** at the restaurant.
Neu:	Chris wanted to know Mika's favorite **fruitssweet** / **safetyarea** by next week.
Inf:	When the accident happened, the staff guided the audience to a **safetyarea** / **fruitssweet** in a hurry.
Neu:	Dan was looking for the nearest **safetyarea** / **fruitssweet** from his house.

(10) babyhelper / floorcover

Inf:	Lisa had worked at a hospital as a **babyhelper** / **floorcover** before she got married.
Neu:	Julia wanted to be a **babyhelper** / **floorcover** when she was young.
Inf:	Bruce's house was dirty so he started cleaning the **floorcover** / **babyhelper** in the living room.
Neu:	The shop began selling **floorcover** / **babyhelper** from this year.

(11) walkroad / breadcake

Inf:	When the runner went into a busy street, he passed a **walkroad** / **breadcake** on his left side.
Neu:	Jim didn't know the place had become a **walkroad** / **breadcake** while he was away.
Inf:	The chef put butter on the **breadcake** / **walkroad** for the guests.
Neu:	Every time Stephanie ate **breadcake** / **walkroad** she was happy.

(12) sealine / hoursewear

Inf:	There was a large shark so nobody went near the **sealine** / **housewear** on that day.
Neu:	Angela saw someone running near the **sealine** / **housewear** on that day.
Inf:	After a hard day's work, Judy went to bed without changing to her **housewear** / **sealine** on that night.
Neu:	When moving into a new house, Alice didn't bring her **housewear** / **sealine** that she no longer liked.

(13) shouesmark / sunclimb

Inf:	When Nichol got lost in the forest, she found a **shoesmark** / **sunclimb** but it wasn't helpful.
Neu:	When Shelly went out, she found a **shoesmark** / **sunclimb** in front of her house.
Inf:	Tommy got up early to see the beautiful **sunclimb** / **shoesmark** with his friends.
Neu:	Robert checked the time of the **sunclimb** / **shoesmark** in the newspaper.

(14) flymachine / nightjob

Inf:	It was Jeremy's first time travelling by **flymachine** / **nightjob** without his parents.
Neu:	From his room, Tony could see a **flymachine** / **nightjob** in the distance.
Inf:	To get more money, Adam also did a **nightjob** / **flymachine** for some years.
Neu:	Dennis was asked to do a **nightjob** / **flymachine** on that day.

(15) lockerbox / homecity

Inf:	When Emi left the place, she put her wallet into a **lockerbox** / **homecity** from her bag.
Neu:	Cameron had never used a **lockerbox** / **homecity** in his life.
Inf:	When Betty became a university student, she left her **homecity** / **lockerbox** for the first time.
Neu:	Jack used the Internet to learn about his **homecity** / **lockerbox** in the computer room.

(16) pricebattle / businessbag

Inf:	The shop closed after the hard **pricebattle** / **businessbag** in the area.
Neu:	Peter studied the reason for the **pricebattle** / **businessbag** on the Internet.
Inf:	After arriving at the office, he realized he had left his **businessbag** / **pricebattle** on the train.
Neu:	When William was cleaning the house, he found a priceba **businessbag** / **pricebattle** ttle but it wasn't his.

Filler sentences (Experiment 1)

(1) Silvia decided that she wanted to be a police officer.
CQ. Did Silvia decide to be a teacher? (No)
(2) Marsha has six cats and two dogs to play with.
CQ. Does Marsha have cats? (Yes)
(3) Fred just got a large bonus from the company.
CQ. Did Fred get a new car? (No)
(4) Wendy was always jumping on the bed in her room.
CQ. Did Wendy jump on her bed? (Yes)

(5) Griffin did not trust the bank at all with his money.
CQ. Did Griffin trust the bank? (No)
(6) The men went bowling to release stress after work.
CQ. Did the men go bowling after work? (Yes)
(7) It was raining outside, so Paul put on a jacket and ran to his house.
CQ. Did Paul go to the station? (No)
(8) The family's cat went outside and was attacked by a dog today.
CQ. Did a dog attack the cat? (Yes)
(9) Tom just graduated from a famous university in the city.
CQ. Did Tom graduate from a high school? (No)
(10) Adam didn't correct the mistake in his report.
CQ. Was there a mistake in the report? (Yes)
(11) John couldn't find his dog when he came back.
CQ. Could John see his dog when he returned? (No)
(12) The fire damaged a lot of books in the library.
CQ. Was there a fire in the library? (Yes)
(13) The child was watching some of the rabbits in the room.
CQ. Was the child chasing chickens? (No)
(14) Luke went to the supermarket and bought an apple.
CQ. Did Luke buy an apple? (Yes)
(15) The group decided to stop some of their projects at the meeting.
CQ. Were they going to continue all the projects? (No)
(16) Helen sent the letter without a stamp so it didn't arrive.
CQ. Was there any mistake with Helen's letter? (Yes)
(17) The visitor to the museum took some rare coins from the shelf.
CQ. Were the coins on the table? (No)
(18) The door to the office was left unlocked by the cleaning service.
CQ. Did someone forget to lock the door? (Yes)
(19) The song was written by a famous American artist.
CQ. Was the song produced by a Chinese person? (No)
(20) Charles was happy when he saw the present on the table.
CQ. Was Charles happy to find the present? (Yes)
(21) Lisa's husband bought a chair at the second-hand shop.
CQ. Did Lisa's husband make the chair? (No)
(22) The coffee shop was famous in that area for its new menu.
CQ. Was the coffee shop famous? (Yes)
(23) Robert collected baseball cards and comic books when he was a child.
(24) The policeman stopped at the bar after work to get a drink.
(25) The gift from the boys was a box of chocolates.
(26) The old town was a popular place for western movies.
(27) Jason had a problem controlling his feelings.
(28) Santa Claus is very important to children in the world.
(29) Ken looked up and saw stars shining in the sky above the mountain top.
(30) They were surprised by the noise from the next room.
(31) Jordan was asked to leave the door open after work
(32) The teacher found that some of his students were always missing.

Appendix 4

Materials Used in Experiment 2

Practice passages

(1)

S1:	Luke went to the supermarket and bought some vegetables.
S2:	He realized that he forgot to buy salad dressing on his way home.
CQ:	What did Luke forget to buy?

Note. The answer for the comprehension question is underlined.

(2)

S1:	John's grandmother refused to buy a computer.
S2:	She has lived without one for eighty years.
CQ:	What didn't John's grandmother buy?

(3)

S1:	The company decided to stop some global projects at the meeting.
S2:	They thought they should focus more on local customers.
CQ:	What did the company decide to stop?

Experimental passages

(1) goundshake /designbook

S1	Inf:	Sara suddenly woke up because she felt a **groundshake** / **designbook** that night.	The student bought a **designbook** / **groundshake** at the art shop.
	Neu:	Sara felt a **groundshake** / **designbook** that night.	The student bought a **designbook** / **groundshake** at the shop.
S2:		The next day, she found that some buildings were damaged.	She was looking forward to drawing pictures with it.
CQ:		What did Sara feel?	What did the student buy?

Note. The first and second target words (boldfaced) served as transparent and opaque target words, respectively. S1 = first sentence, S2 = second sentence, Inf = informative context, and Neu = neutral context.

(2) lunchhour / waterwear

S1	Inf:	Kevin had to hold a meeting during **lunchhour** / **waterwear** because he was very busy.	During the summer, the sport shop sold **waterwear** / **lunchhour** at a low price.
	Neu:	Kevin held a meeting during **lunchhour** / **waterwear** on that day.	The shop sold **waterwear** / **lunchhour** to make greater profits.
S2:		He asked the employees to bring some food during the meeting.	Many people bought it because a new pool had opened.
CQ:		When did Kevin hold a meeting?	What did the shop sell?

(3) trainroad / paintstick

S1	Inf:	For the convenience of commuters, a **trainroad** / **paintstick** was built in the area.	Emi went to an art shop and bought a **paintstick** / **trainroad** that she wanted.
	Neu:	Construction took place to build a **trainroad** / **paintstick** in the area.	Emi bought a **paintstick** / **trainroad** that she had seen on TV before.
S2:		It allowed people to move to the city center easily.	She decided on one that could draw thin lines.
CQ:		What was built in the area?	What did Emi buy?

(4) gunball / hotelstaff

S1	Inf:	Robert bought a **gunball** / **hotelstaff** to protect himself from strangers.	Yumi got his room key and gave her bags to a **hotelstaff** / **gunball**.
	Neu:	Robert bought a **gunball** / **hotelstaff** and went home directly.	Yumi gave her bags to a **hotelstaff** / **gunball** and left the place.
S2:		It was dangerous, so he put it in a safe space.	He kindly offered to bring them to his room.
CQ:		What did Robert buy?	To whom did Yumi give his bag?

(5) lightpole / meatblock

S1	Inf:	The city built a **lightpole** / **meatblock** along the street for the safety of the citizens.	Jessica bought a **meatblock** / **lightpole** at the supermarket for a barbecue party.
	Neu:	The city built a **lightpole** / **meatblock** along the street.	On her way home, Jessica bought a **meatblock** / **lightpole** for dinner.
S2:		It reduced traffic accidents at night in the area.	She bought one that looked both juicy and delicious.
CQ:		What did the city build?	What did Jessica buy?

(6) guardperson / handknife

S1	Inf:	They decided to hire a **guardperson** / **handknife** to protect their important guests.	The boy was cooking outside and used a **handknife / guardperson.**
	Neu:	They decided to hire a **guardperson** / **handknife** for the meeting.	The boy's mother taught him how to use a **handknife / guardperson**
S2:		They wanted someone who was both strong and careful.	He was good at cutting food and enjoyed the meal.
CQ:		Whom did they decided to hire?	What did the boy use?

(7) campspot / morninglife

S1	Inf:	The group went to a **campspot** / **morninglife** and did some outdoor activities.	Steve followed his doctor's advice and started **morninglife** / **campspot** for his health.
	Neu:	Tom and Cathy went to a **campspot** / **morninglife** on that day.	Steve started **morninglife** / **campspot** last week.
S2:		They enjoyed the fresh air and beautiful nature there.	Soon he found it difficult to change his life style.
CQ:		Where did they go?	What did Steve start?

(8) moneyticket / trainstop

S1	Inf:	Jill received a **moneyticket** / **trainstop** from the shop for the inconvenient service she received.	Matt bought a house close to the **trainstop** / **moneyticket** because he travelled often.
	Neu:	Jill received a **moneyticket** / **trainstop** on that day.	Matt bought a house close to the **trainstop** / **moneyticket** last year.
S2:		She was lucky and planned what to buy with it.	It was expensive, but it saved him travel time.
CQ:		What did Jill receive?	What was near Matt's house?

(9) fruitssweet /safetyarea

S1	Inf:	For dessert, Jane ate a **fruitssweet** / **safetyarea** at the restaurant.	When the accident happened, the customers looked for a **safetyarea** / **fruitssweet**.
	Neu:	Jane ate a **fruitssweet** / **safetyarea** at the restaurant.	The customers at the supermarket looked for a **safetyarea** / **fruitssweet**.
S2:		She liked its sweet taste and its beautiful decoration.	Soon the staff guided them, and no one was injured.
CQ:		What did Jane eat?	What did the customers look for?

(10) babyhelper / floorcover

S1	Inf:	Lisa liked small children and worked as a **babyhelper** / **floorcover** before she got married.	The furniture shop imported a **floorcover** / **babyhelper** last month.
	Neu:	Lisa worked as a **babyhelper** / **floorcover** before she got married.	The shop imported a **floorcover** / **babyhelper** last month.
S2:		The experience was helpful when she had her own babies.	It was comfortable to both sit and sleep on.
CQ:		What was Lisa's job?	What did the shop import?

(11) walkroad / breadcake

S1	Inf:	When the event was held on the street, the staff told the people to pass a **walkroad** / **breadcake**.	The chef cooked a **breadcake** / **walkroad** and put butter and fruit on it.
	Neu:	The staff told the people to pass a **walkroad** / **breadcake**.	The chef cooked a **breadcake** / **walkroad** and gave it to customers.
S2:		Some people walking on the street got severely injured.	It was soft, and many children liked its sweet taste.
CQ:		Where were the people told to pass?	What did the chef cook?

(12) sealine / housewear

S1	Inf:	When flying above the sea, Angela saw a **sealine** / **housewear** clearly from a plane.	At the clothing store, Judy bought a **hosewear** / **sealine** for herself.
	Neu:	When Angela was on a plane, she saw a **sealine** / **housewear** clearly.	Judy went to a shopping center and bought **housewear** / **sealine** for herself.
S2:		It was beautiful, and she wanted to go there and swim.	It fit her body and was comfortable to move around in.
CQ:		What did Angela see?	What did Judy buy?

(13) shoesmark / sunclimb

S1	Inf:	When walking in the forest, Nicole found a **shoesmark** / **sunclimb** in the snow.	Tony got up early and went out to see the **sunclimb** / **shoesmark** on that day.
	Neu:	After walking for an hour, Nicole found a **shoesmark** / **sunclimb** on the ground.	Tony saw a **sunclimb** / **shoesmark** in the morning.
S2:		She thought someone had passed there a while ago.	The view was beautiful, and he decided to see it again.
CQ:		What did Angela see?	What did Tony see?

(14) flymachine / nightjob

S1	Inf:	Last month, Betty used a **flymachine** / **nightjob** to travel for the first time.	To get more money, Adam had to do a **nightjob** / **flymachine** for some years.
	Neu:	Last month, Betty used a **flymachine** / **nightjob** for the first time.	Adam had to do a **nightjob** / **flymachine** for some years.
S2:		She could not see outside of it while it was moving.	He was able to save money, but it damaged his health.
CQ:		What did Betty use?	What did Adam have to do?

(15) lockerbox / homecity

S1	Inf:	The staff asked Emi to use a **lockerbox** / **homecity** when leaving important things.	Last year, Shelly left her **homecity** / **lockerox** for her university life.
	Neu:	Emi was asked to use a **lockerbox** / **homecity** by the staff.	Last year, Shelly had to leave her **homecity** / **lockerox** for some reasons.
S2:		She put her wallet and mobile phone in it.	A month later, she missed her family and returned there.
CQ:		What was Emi asked to use?	Where did Shelly leave?

(16) pricebattle / businessbag

S1	Inf:	The shop reduced its production costs because it experienced a **pricebattle** / **businessbag** in the area.	William realized that he had left his **businessbag** / **pricebattle** when he arrived at the office.
	Neu:	The shop experienced a **pricebattle** / **businessbag** in the area.	William realized that he had left his **businessbag** / **pricebattle** somewhere.
S2:		It was so severe that the shop quickly shut down.	He was upset because his computer was in it.
CQ:		What did the shop experience?	What did William leave?

Filler passages

(1)

S1:	Helen received a letter from her friend in France.
S2:	She was happy that her friend would visit her a month later.
CQ:	What did Helen receive from her friend?

Note. The answer for the comprehension question is underlined.

(2)

S1:	Dale liked playing video games.
S2:	He spent most of his free time on the computer.
CQ:	What did Dale like to play?

(3)

S1:	Sandy watched an action movie with her friend.
S2:	It was boring, and she wanted her money back.
CQ:	What did Sandy watch?

(4)

S1:	Julia liked to walk her dog in the park.
S2:	There was a lot of space for animals to play there.
CQ:	Where did Julia like to walk her dog?

(5)

S1:	Robert collected baseball cards as a child.
S2:	He had a chance to sell them twenty years later.
CQ:	What did Robert collect?

(6)

S1:	Frank looked up and saw stars in the sky.
S2:	They were so beautiful that he stopped there for an hour.
CQ:	What did Frank see?

(7)

S1:	The policemen stopped at the bar after work to get a drink.
S2:	They heard it was the best bar in town.
CQ:	Where did the policemen go?

(8)

S1:	It was raining outside, so Paul put on a jacket and ran in the rain.
S2:	It kept him warm in the freezing conditions.
CQ:	What did Paul put on?

(9)

S1:	Phillipe did not trust the bank at all.
S2:	He felt like they would steal money from him.
CQ:	What doesn't Phillipe trust?

(10)

S1:	John couldn’t find the pen case he had borrowed from his friend.
S2:	He looked for it around the house and found it under the bed.
CQ:	What was John looking for?

(11)

S1:	The visitor to the museum saw some clothes in the display.
S2:	He knew that they were used by a famous emperor.
CQ:	What did the visitor see?

(12)

S1:	The door to the office was left unlocked by the cleaning staff.
S2:	The manager got angry but found no one had entered the building.
CQ:	Who forgot to lock the door?

(13)

S1:	The book was written by an athlete.
S2:	It sold well and was translated into many languages.
CQ:	Who wrote the book?

(14)

S1:	Charles was happy when he saw the present on the table.
S2:	He opened it and took it to his parents.
CQ:	Where did Charles find the present?

(15)

S1:	The fire damaged some history books in the library.
S2:	The staff was shocked because they were very rare.
CQ:	What were damaged?

(16)

S1:	Jason had trouble controlling his feelings.
S2:	He often broke his computers when he got angry.
CQ:	What did Jason often break?

(17)

S1:	The old town was a popular place for movies.
S2:	Tourists often visited there to see the site.
CQ:	Who often visited the town?

(18)

S1:	The university campus store was shut down last week.
S2:	The store owner had been selling cigarettes to children.
CQ:	What did the owner sell?

(19)

S1:	Tom graduated from a famous university.
S2:	He was looking for a job at a hospital.
CQ:	Where did Tom want to work?

(20)

S1:	George recently found out that his wife was expecting a baby.
S2:	He soon called his parents to tell them the good news.
CQ:	Whom did George call?

(21)

S1:	Kate's cat was bitten by a dog on that day.
S2:	She went to an animal hospital soon.
CQ:	Where did Kate go with her cat?

(22)

S1:	Nick had a house with a backyard.
S2:	He was planning to make a home garden there.
CQ:	What was Nick planning to make?

(23)

S1:	Susan was interested in watching car races.
S2:	She wanted a driver's license as soon as possible.
CQ:	What did Susan want?

(24)

S1:	Marsha had six cats and two dogs to play with.
S2:	She often took her pets to her company.
CQ:	Where does Marsha take her pets?

(25)

S1:	Fletcher just got a large bonus from the firm.
S2:	He decided to buy his family a sofa.
CQ:	What did Fletcher decide to buy?

(26)

S1:	Adam had no idea about what to write for the essay.
S2:	He decided to go to the teachers' office and asked for advice.
CQ:	Where did Adam decide to go?

(27)

S1:	The fire damaged some history books in the library.
S2:	The staff was shocked because they were very rare.
CQ:	What were damaged?

(28)

S1:	The child was watching some of the rabbits in the room.
S2:	He was planning to make a fence for them.
CQ:	What was the child planning to make?

(29)

S1:	Lisa bought a chair at a home center.
S2:	She planned to put it in the kitchen.
CQ:	Where did Lisa plan to use the chair?

(30)

S1:	The boy got a lot of birthday presents at the school.
S2:	He put them in his bag and was looking forward to opening them.
CQ:	Where did the boy put his presents?

(31)

S1:	Dennis was worried by the noise from the next room that night.
S2:	He couldn't concentrate and went to the library.
CQ:	Where did Dennis go?

(32)

S1:	Jordan was asked not to use the computer room.
S2:	Computer maintenance was being carried out there.
CQ:	What was being carried out?

Appendix 5

Individual Pattern of Literal Interpretations for Opaque Targets in Experiment 3

	The number of encounters of opaque words																
Participant	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	12th	13th	14th	15th	16th	Total
1	PM																1
2		NC	MC														2
3				NC													1
4		PM									NC						2
5	PM																1
6																	0
7	LE								MC				EC		EC		4
8						MC											1
9																	0
10	PM																1
11		NC		EC										LE			3
12																	0
Total	4	3	1	2	0	1	0	0	1	0	1	0	1	1	1	0	16

Note. NC = Negligence of Context, LE = Lack of evaluation, PM = Persistent reliance of morphemes, MC = Modification of context, EC = Elaboration of context.

Appendix 6

Summary of English Proficiency Levels of the Present Participants

			EIKEN Test					TOEIC L&R				TOEFL ITP				Reading proficiency test		
Study	Experiment	*N*	4	3	pre-2	2	pre-1	*n*	*M* (SD)	*Min*	*Max*	*n*	*M* (*SD*)	*Min*	*Max*	*M* (*SD*)	*Min*	*Max*
Study 1	Experiment 1	30	0	6	1	6	2	5	706.00 (54.93)	620	765	3	490.33 (22.03)	469	513	15.20 (2.64)	11	21
	Experiment 2	24	0	1	5	6	4	13	751.92 (126.59)	495	905	1	580.00 (0.00)	580	580	16.00 (3.38)	12	24
	Experiment 3	12	0	0	2	5	0	3	731.67 (98.52)	630	865	2	519.00 (34.19)	490	567	13.75 (2.92)	11	20
Study 2	Experiment 4	21	0	9	4	0	0	0	NA	NA	NA	0	NA	NA	NA	9.71 (2.12)	6	14
	Experiment 5 (Upper)	37	2	10	3	0	0	1	250 (0.00)	250	250	0	NA	NA	NA	10.73 (1.54)	9	13
	Experiment 5 (Middle)	41	1	10	5	0	0	0	NA	NA	NA	0	NA	NA	NA	6.93 (0.72)	6	8
	Experiment 5 (Lower)	41	2	11	0	0	0	0	NA	NA	NA	0	NA	NA	NA	4.00 (1.02)	1	5
	Experiment 5 (Total)	119	5	31	8	0	0	1	250 (0.00)	250	250	0	NA	NA	NA	7.10 (2.96)	1	13

Note. NA = not applicable. The reading proficiency test was conducted in this study. The maximum possible test score for Experiments 1, 2, and 3 was 24 (pre-2nd: $k = 4$; 2nd: $k = 15$; pre-1st: $k = 5$), and that for Experiments 4 and 5 was 20 (3rd: $k = 8$; pre-2nd: $k = 8$; 2nd: $k = 4$). The time allocated for the test was 20 minutes.