Construction of Globally Coherent Mental Representations of Texts in Japanese EFL Learners' Reading Comprehension

A Dissertation

Submitted to the University of Tsukuba

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Linguistics

Yukino KIMURA

2015

Abstract of the Dissertation

Construction of Globally Coherent Mental Representations of Texts

in Japanese EFL Learners' Reading Comprehension

by

Yukino KIMURA

Reading comprehension is an activity that requires a reader to understand not only explicitly stated individual words and sentences, but also the broader message conveyed by a writer. However, a writer's message is often stated implicitly rather than explicitly; therefore, readers must go beyond the explicit information provided and use contextual information or background knowledge to uncover a writer's intended meaning.

Although specific details of theoretical models of text comprehension differ (e.g., Gernsbacher, 1990; Kintsch, 1998; van den Broek, Risden, Fletcher, & Thurlow, 1996; van Dijk & Kintsch, 1983), they commonly define reading comprehension as the construction of a meaningful and coherent mental representation of a text. To build a coherent mental representation, readers are required to generate *inference*. When the reader generates inference to construct the global message or point of a text, deeper comprehension is achieved (e.g., Graesser, Pomeroy, & Craig, 2002; Graesser, Singer, & Trabasso, 1994).

The current research aimed to investigate Japanese English as a foreign language (EFL) learners' construction of globally coherent representations of texts, focusing on two types of inference: *thematic inference*, which is a point, message, or moral in narrative comprehension, and *superordinate inference*, which is the relationship between a sequence of statements subsumed under the overall text information. To date, previous studies have empirically

investigated how readers constructed globally coherent representations through inference generation in first language (L1) reading (e.g., Brown, Day, & Jones, 1983; Dorfman & Brewer, 1994; Kurtz & Schober, 2001; Seifert, McKoon, Abelson, & Ratcliff, 1986; Ritchey, 2011). In contrast, most research on second language (L2) and EFL reading examined several types of inference concurrently (e.g., Horiba, 1996; Muramoto, 2000; Yoshida, 2003) without specific focus on inference that contributes to global coherence of a text.

Considering the importance of understanding the broader messages of a text, investigating whether and how Japanese EFL learners construct globally coherent mental representations of texts will have both theoretical and educational implications. The current research was composed of the following six experiments, conducted in order to examine the construction of globally coherent mental representations among Japanese EFL learners.

Study 1 of this dissertation included three experiments (Experiments 1–3) designed to examine thematic inference generation in narrative reading. Experiment 1 examined two questions: whether Japanese EFL learners generate thematic inference to understand implicit themes in narrative texts, and whether the generation of thematic inference differs from other types of inference. This experiment manipulated the explicitness of thematic statements in narrative passages and used an inference verification task requiring Japanese university students to evaluate whether statements could be reasonably inferred from the passages. The results indicated that learners evaluated implicit themes with lower validity than explicit themes, suggesting that understanding implicit themes through inference generation can be difficult for them. In addition, thematic inference was generated more than emotional inference, but at a comparable level of goal, action, and state inferences.

Experiment 2 investigated whether prompting Japanese EFL readers to attend to the theme of a narrative passage with task instructions facilitated thematic inference generation and improved text comprehension. The Task condition (i.e., learners instructed to read passages in

order to understand the theme conveyed by the writer) was compared with the Control condition (i.e., learners instructed to read passages for comprehension) on performance of a thematic inference task and a written recall task. The results demonstrated that task instructions effectively facilitated the learners' thematic inference generation and improved text comprehension.

Experiment 3 investigated whether task instructions aimed at thematic inference generation changed EFL learners' reading goals or altered EFL learners' cognitive processes. Experiment 3 had the same task instructions as Experiment 2, with the addition of a think-aloud task and a brief questionnaire. The results indicated that task instructions altered the learners' reading goals and reduced allocation of cognitive resources to lower-level linguistic processes, such as analyzing individual words and sentences. Further analysis demonstrated that while both proficient and less proficient learners tried to employ different cognitive strategies according to the task instructions, higher-proficiency learners more flexibly changed their cognitive processes according to their reading goals than lower-proficiency learners.

Study 2 included three experiments (Experiments 4–6) designed to investigate superordinate inference generation in expository reading. Experiment 4 attempted to answer two questions: whether Japanese EFL learners could generate inference to understand implicit superordinate propositions in expository texts, and whether the presence of superordinate propositions affects a learner's mental representation of expository text. This experiment manipulated the explicitness of superordinate propositions in expository texts followed by administration of an inference verification task. Results suggested that while superordinate inferences were likely to be generated spontaneously, without specific task instruction, the experiment should be replicated with different methodology. In addition, when the superordinate proposition was implicit, learners were less likely to suppress activation of information inconsistent with global text representations.

Experiment 5 examined the effects of task instructions on superordinate inference generation and text comprehension. Experiment 5 assessed the difference between the Task condition (i.e., learners instructed to read expository texts in order to understand the message conveyed by the writer) and the Control condition (i.e., learners instructed to read the texts for comprehension). Results showed that success of superordinate inference was not related to type of task instructions, but was related to L2 reading proficiency. Learners were likely to construct narrow representations based on the main idea of each paragraph, rather than the overall text. These findings suggest that learners had difficulty integrating and constructing information distributed across paragraphs, regardless of the task instructions. Furthermore, task instructions demonstrated no effect on the written recall task.

Experiment 6 explored the effects of integration task, which prompted EFL learners to integrate information across paragraphs, on superordinate inference generation, cognitive processes during reading, and text comprehension. Learners were instructed to identify the writer's intended message at the conclusion of each paragraph. Results demonstrated that the integration task effectively facilitated superordinate inference generation and text comprehension. Conversely, results of the think-aloud task demonstrated that cognitive processes during reading did not change according to the reading task, due to over allocation of resources to lower-level processing in expository reading. Additional analysis demonstrated that reading task did not directly change a learner's processes during reading, whereas they did influence goals for reading, which resulted in improved text comprehension.

The main findings of the present study can be summarized by the following three points: (a) thematic and superordinate inference are strategic rather than automatic processes in EFL text comprehension; (b) expository reading is more likely to be influenced by a learner's L2 reading proficiency than task instructions, whereas narrative reading is more likely to be influenced by task instructions than L2 reading proficiency; and (c) EFL learners attempt to

alter their cognitive processes according to reading goals prompted by task instructions for both narrative and expository reading.

Based on these findings, some pedagogical implications for the construction of globally coherent mental representations for EFL readers were provided. The present study is significant in demonstrating not only whether learners can generate global inference but also how educators can solve the difficulties encountered in an EFL learner's reading processes. However, to identify additional theoretical and pedagogical implications for L2/EFL reading, further research investigating reading comprehension at global levels is needed.

Acknowledgements

I owe a debt of gratitude to many people for their invaluable and beneficial assistance through the course of my studies at the Graduate School of Humanities and Social Sciences at the University of Tsukuba. Although it would be difficult to acknowledge everyone who has assisted me in writing this dissertation, I would especially like to thank the following people.

First of all, I would like to express the deepest appreciation to my academic supervisor, Professor Yuji Ushiro at the University of Tsukuba, who awakened my interest in English education and in second language learners' reading. What I learned in his lecture and seminar inspired me to study this field more deeply. He gives me beneficial advice from both theoretical and educational viewpoints to improve this dissertation. In addition, his warm encouragement always motivates me to engage in my research.

I would also like to express my appreciation to Professor Hirosada Iwasaki at the University of Tsukuba. He gave me helpful comments on experimental design from the view of corpus linguistics. Moreover, what I learned from his lecture about academic writing greatly helps me to write this dissertation. Professor Akira Kubota at the University of Tsukuba also gave various suggestions from the perspectives of second language acquisition. I appreciate his insightful comments on my research. I am also deeply grateful to my external supervisor, Professor Yukie Horiba at Kanda University of International Studies. Her meticulous comments greatly improved an earlier version of this dissertation.

Many thanks are also due to my senior associates, especially Dr. Haruka Shimizu, Dr. Shingo Nahatame, Dr. Yusuke Hasegawa, and my junior associates, especially Ms. Natsumi Tanaka, Mr. Yoshinobu Mori, and Mr. Masaya Hosoda for their beneficial support throughout the course of my research. I would also like to express my appreciation to my colleague Mr. Akira Hamada. He always gives me insightful comments, and I respect his effort and attitude

toward his research.

Thanks also to 235 undergraduate and graduate students for participating in the experiments. Without their kind cooperation, I would not have conducted my experimental studies.

Finally, I would like to express my deep gratitude to my parents for their understanding, encouragement, and warm support. They always believe in me and never stopped encouraging me.

Yukino KIMURA

Table of Contents

List of Tables

List of Figures

List of Appendices

Chapter 1
Introduction

1.1 Background of the Current Research

Reading is an activity where readers and writers interact to construct meaning; that is, the writers of texts intentionally write to convey the message, and readers attempt to understand the writer's intended meaning. However, writers do not always state their points, messages, and claims explicitly in the texts. In that case, readers need to go beyond the explicit information provided and use the context or their background knowledge to uncover what the writer intended to convey. On the other hand, especially in second language (L2) or English as a foreign language (EFL) reading, while readers can comprehend individual words and sentences in the text, they sometimes do not grasp the overall message of the text. It is said that readers "do not see the forest for the trees" in reading comprehension.

In the field of research in first language (L1) reading, many researchers have tried to investigate how readers construct meaningful and coherent representation through inference generation (Graesser, Singer, & Trabasso, 1994; McKoon & Ratcliff, 1992; van Dijk & Kintsch, 1983). Among various types of inference, the current research focused on *thematic inference*, which is a point, message, or moral in narrative comprehension, and *superordinate inference*, which is the relation between a sequence of statements subsuming the overall text information, in expository reading. Since these inferences contribute to building globally coherent mental representations of the texts, they are necessary for readers to understand implicit messages in the texts. Look at the following example of the summary of a well-known story, "The Tortoise and the Hare."

A tortoise and hare lived in the woods. One day the hare challenged the tortoise to a race.

> The tortoise plodded along at his usual slow pace, while the hare, sure of his swift speed, took a nap. Eventually, the tortoise passed the sleeping hare and won the race.
>
> (Dorfman & Brewer, 1994, p. 107)

In this short passage, readers can generate inference such as "slow and steady wins the race," based on the contextual information and their background knowledge, although this message is not explicitly stated. If a reader fails to construct a message that satisfies various constraints in the explicit text, then the reader has somehow failed to build a cognitive representation that is globally coherent to the text.

To date, previous studies have empirically investigated how readers construct globally coherent representations through inference generation in first language (L1) reading (e.g., Brown, Day, & Jones, 1983; Dorfman & Brewer, 1994; Kurtz & Schober, 2001; Ritchey, 2011; Seifert, McKoon, Abelson, & Ratcliff, 1986). Despite the importance of building global coherence of the text, while some L2 and EFL reading research have examined what types of inferences can be generated in reading comprehension (e.g., Horiba, 1996; Muramoto, 2000; Yoshida, 2003) they did not pay particular attention to inference that contributes to global coherence of the text. Due to this, this study aims to investigate whether and how Japanese EFL learners construct globally coherent mental representations of narrative and expository texts, focusing on thematic inference and superordinate inference. The present study will be a first step to demonstrate the processes in which Japanese EFL learners construct globally coherent mental representations of the overall texts and it will suggest the effective methods of instruction for struggling EFL readers who "do not see the forest for the trees" in reading comprehension.

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).

In Chapter 2, theories and models of reading comprehension (e.g., mental representation of the text, the construction-integration model, the structure-building framework, the landscape model), narrative and expository text comprehension, establishing text coherence through inference generation, and effects of task instructions on reading comprehension are reviewed. The findings and limitations of the previous studies are finally summarized in this chapter.

In order to examine the construction of globally coherent mental representations among Japanese EFL learners, the current research conducted a total of six experiments. First, Study 1 of this dissertation conducted three experiments (Experiments 1 to 3) to examine thematic inference generation in narrative reading. Study 2 conducted three experiments (Experiments 4 to 6) to investigate superordinate inference generation in expository reading. The overview of the experimental studies is illustrated in Figure 1.1.

In Chapter 3, Study 1 explored construction of globally coherent mental representations of narrative texts, focusing on thematic inference generation. Experiment 1 examined whether Japanese EFL learners understand implicit themes in narrative texts by generating thematic inference and whether the generation of thematic inference differ from other types of inference. In Experiment 2, whether task instructions to comprehend the theme of the narrative passages facilitate thematic inference generation and text comprehension among Japanese EFL readers were investigated. Experiment 3 investigated whether the task instructions aimed at thematic inference generation alter EFL learners' reading goals, cognitive processes, and text comprehension.

In Chapter 4, Study 2 investigated construction of globally coherent mental

representations of expository texts, focusing on superordinate inference generation. Experiment 4 was conducted to examine whether Japanese EFL learners can understand implicit superordinate proposition of expository texts through inference generation and whether the presence of the superordinate proposition affect the learners' mental representations of expository texts. In Experiment 5, the effects of task instructions on superordinate inference generation and text comprehension were examined. Experiment 6 explored the effects of integration task which induced EFL learners to integrate information distributed among paragraphs on superordinate inference generation, cognitive processes during reading, and text comprehension.

Chapter 5 generally discusses the results of the six experimental studies, and Chapter 6 concludes the construction of globally coherent mental representations of texts in Japanese EFL learners' reading comprehension. Finally, this dissertation summarizes the limitations and suggestions for future research, and introduces some pedagogical implications for EFL reading instructions.

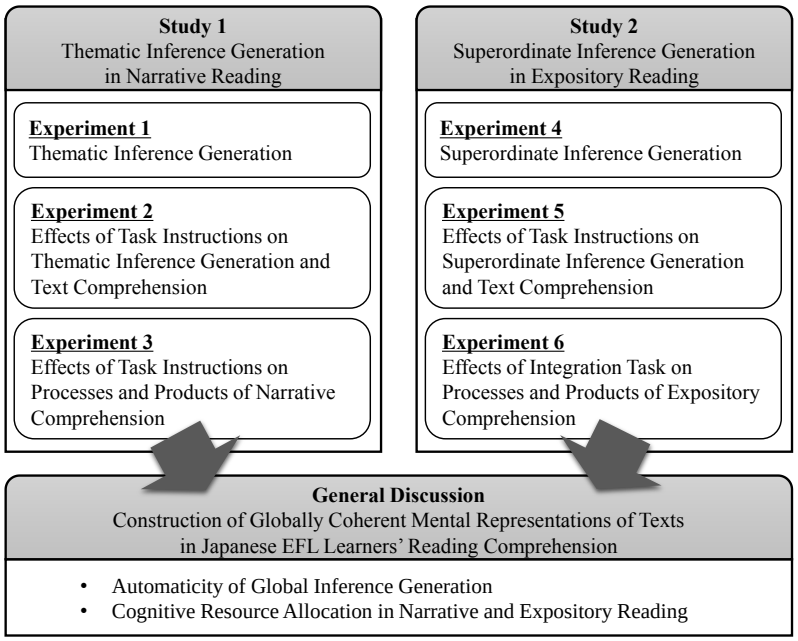

Figure 1.1. Overview of the six experiments in the present study.

Chapter 2

Review of Related Literature

2.1 Reading Comprehension

2.1.1 Theories and Models of Reading Comprehension

Early research on text comprehension was interested in how text information is represented in readers' memory (e.g., Johnson-Laird, 1983; van Dijk & Kintsch, 1983). Readers comprehend the text by building and remembering the text representation in their mind. The representation is called *mental representation* and the degree of text comprehension depends on how elaborate and coherent their representations are. Mental representation is based on what we read and remember, so it involves not only text elements, but also knowledge elements that readers have (van Dijk & Kintsch, 1983).

van Dijk and Kintsch classified the mental representation into three phases as follows: (a) *surface memory*, (b) *propositional textbase*, and (c) *the situation model*. These three phases have been clearly distinguished by empirical studies (e.g., Kintsch, Welsch, Schmalhofer, & Zimny, 1990; Mulder & Sanders, 2012). First, surface memory is the representation for the text form featured by the words and phrases as they are used in the original text. Although this surface structure may be initially available in memory, this level of representation has been shown to decay rapidly (Kintsch et al., 1990). The second phase is propositional textbase that is the memory for individual propositions in the text and captures the meaning conveyed by the text. Proposition is defined as the smallest unit that refers to states, phenomena and actions (Kintsch, 1998). In this phase, the meaning of the text is more important than in surface memory. The third phase is called the situation model and has the deepest levels of comprehension. The situation model is the overall image of the text and is constructed by integrating readers' prior knowledge with the text information; therefore, the richness of mental representation depends

on how much inferences they generate by activating their background knowledge.

The situation model regards mental representations as multidimensional that include five dimensions: time, space, causation, motivation, and protagonist (e.g., Zwaan, Langston, & Graesser, 1995; Zwaan & Radvansky, 1998). According to the *event-indexing model*, readers organize each event in their mental representations based on these dimensions. For example, when an event was continuous with the other event (e.g., the same protagonist was in the two events), these events were strongly connected in readers' mental representations (e.g., Rinck & Weber, 2003; Zwaan, 1996). For example, Zwaan demonstrated that sentence reading times were longer when there was chronological distance between two story events (e.g., *an hour later*) rather than no chronological distance (e.g., *a moment later*). Rinck and Weber also showed that reading times increased for protagonist and temporal shifts, but the effect of spatial shifts was less significant. Thus, these processing loads on situational shifts supported the notion that readers monitor and represent these multiple dimensions at the same time to construct coherent mental representations. The extent which readers engage in maintaining coherence of mental representations was also dependent on the readers' standards of coherence (see *landscape model* below).

As well as explaining the content of readers' metal representations reviewed above, the researchers also examined the processes underlying text comprehension and proposed theories. In the following sections, three models related to the current research are reviewed: the construction-integration model (Kintsch, 1988, 1998), structure-building framework (Gernsbacher, 1990), and the landscape model (van den Broek, Risden, Fletcher, & Thurlow, 1996). The construction-integration model and structure-building framework both explain text comprehension inclusively while the landscape model aims to integrate hypotheses derived from some models regarding text comprehension.

The construction-integration model (Kintsch, 1988, 1998)

The construction-integration model explains how text information and readers' background knowledge interact to build a mental representation of the text. This model assumes that text comprehension involves the following two steps: *construction* and *integration* phases.

In the construction phase, readers construct a semantic network based on the textual information and related background knowledge. In this phase, readers activate related knowledge based on associative priming rather than strategic processing. Because all concepts and ideas from the text and readers' prior knowledge are initially activated, the mental representations constructed in this phase are incoherent, containing irrelevant information. To eliminate the irrelevant information and to strengthen relevant elements, readers engage in the integration phase. In this phase, text information and activated knowledge that have many connections are integrated in the readers' mental representations, while irrelevant information is deactivated and disappears from the representations.

The results of simulations of data based on this model showed that the human performance of comprehension and learning tasks had a moderate correlation with the construction-integration model, and the model typically predicted off-line memory tasks better than on-line processing tasks (e.g., Kintsch, 1998; Singer & Kintsch, 2001).

The structure building framework (Gernsbacher, 1990)

According to the structure building framework (Gernsbacher, 1990), the goal of comprehension is to construct coherent mental representations called "structure." In reading comprehension, readers activate relevant information from their long-term memories to form a foundation of structure (i.e., foundation laying), and then integrate incoming information with previously introduced ideas to develop the structure (i.e., structure mapping). However, when the incoming information is inconsistent with the current structure the reader needs to build a

new structure (i.e., structure shifting).

Gernsbacher (1990, 1996) explained these three structure building processes on the basis of many phenomena that occur during reading. For example, as for the step of foundation laying, reading time on the first word of a clause or sentence, and the first sentence of a paragraph, becomes longer because the first segments are used to form foundations of mental representations. Cognitive effort is allocated to lay a foundation, which consequently increases the reading time of the first segment. Regarding the step of structure mapping, the reading time on referentially, temporally, and locationally coherent sentences with previous information is faster than on sentences that are less coherent with previous information. As for structure shifting, this process was explained by the phenomena where readers spend more time on words and sentences that change the topic, point of view, locational settings, and temporal settings than information that does not involve such changes.

Although the structure building framework is similar to the construction-integration model in that it explains how coherent mental representations are constructed, they differ as follows: While the construction-integration model assumes that irrelevant information is suppressed automatically in the integration process, the structure-building framework assumes that such suppression is based on strategic process rather than automatic process. Gernsbacher, Varner, and Faust (1990) empirically demonstrated that the differences between skilled and less skilled comprehenders lies in whether they can suppress the irrelevant information efficiently.

Landscape model (van den Broek et al., 1996)

Van den Broek and his colleagues proposed the landscape model by further developing the construction-integration model. This model differs from other models in that (a) it assumes both on-line and off-line processes of comprehension and (b) it considers the readers' *standards of coherence* as the factor that affects the activation of information.

The landscape model captures the dynamic and reciprocal interaction between on-line processes and the off-line products of reading. Although a large number of studies have investigated reading processes and memory representations, few theories explain the relationship between on-line processes and off-line representations. According to the landscape model, reading is a cycling process where propositions or other units fluctuate and activate in each cycle. The sources of activation are various, such as the current information explicitly mentioned in the text, the previous information in the prior cycle, the reinstatement of concepts from the prior text, and background knowledge. These various sources determine the patterns of activation during reading, which creates a landscape of fluctuating activations, with peaks and valleys in activation for individual concepts.

The memory representation of the text and background knowledge can be activated and accessed through automatic *cohort activation* or strategic *coherence based retrieval*. Cohort activation is a process where relevant concepts are activated and connected to each other to form a cohort. This activation is passive and memory-based processing. In contrast, coherence based retrieval is a strategic mechanism where readers retrieve relevant information in order to achieve their reading goals. Whether readers engage in these automatic and strategic processes depends on the readers' standards of coherence.

The reader's standards of coherence is defined as the types and strengths of comprehension that a reader attempts to maintain during reading (van den Broek, Bohn-Gettler, Kendeou, & Carlson, 2010; van den Broek, Lorch, Linderholm, & Gustafson, 2001; van den Broek, Risden, & Husebye-Hartmann, 1995). Van den Broek et al. (2010) summarized the roles of standards of coherence in reading processes as the following four points: First, readers adopt these standards for a particular reading situation implicitly or explicitly. Second, if the activation of concepts by automatic processes is not sufficient for their standards, then readers engage in strategic processes until the levels of comprehension meet those standards. Third,

while some types of coherence (e.g., referential, causal) are adopted in most reading situations, other types of coherence (e.g., spatial) are adopted only when given specific reading goals. Fourth, the standards of coherence are affected by the text, reader, and task characteristics.

Thus, based on the standards, readers determine whether the cohort activation is adequate for their reading goals, or whether strategic processes are necessary. For example, less skilled readers often have lower standards of coherence than skilled readers; therefore, they are likely to be satisfied with less coherent mental representations and not engage in cognitive processes for coherence building. Other empirical studies regarding readers' standards of coherence are reviewed in later section (see section 2.3).

In sum, although theoretical models of text comprehension differ in detail, they are common in that the core of reading comprehension is the construction of meaningful and coherent mental representations of texts by connecting text elements. Moreover, construction of coherent mental representations of texts arises as a result of the various automatic/strategic activations and suppressions of text information and background knowledge. While the construction-integration model regards both activation and suppression as automatic processes, the structure building framework regards the suppression processes as less automatic. The landscape model describes reading comprehension as fluctuating processes where various concepts are activated automatically and strategically in a reading cycle. Finally, each model (especially the landscape model) assumes that reading comprehension is a result of the interaction between texts, readers, and tasks.

2.1.2 Narrative and Expository Text Comprehension

Among various genres of texts, two major categories are narrative texts and expository texts (Graesser & Goodman, 1985). Although defining clear distinctions between these two

types is difficult, the largest difference between narrative and expository texts is its objective. As for narrative texts, they describe a temporal and causal sequence of events; therefore, the purpose of reading narratives is to understand the emotions and points of view of the writers or characters. Narrative texts have a close correspondence to everyday experience, in that people perform actions to achieve goals while obstacles and emotional reactions occur during the event. Therefore, it is comparatively easy for adults to construct situation models for narrative texts (Freedle & Halle, 1979). On the other hand, expository texts are defined as informational texts that explain unfamiliar concepts and relations for the reader; therefore, the main objective of reading expository texts is to learn unfamiliar information (Coté, Goldman, & Saul, 1998; Wolfe & Woodwyk, 2010).

Not only do the objectives of narrative and expository reading differ, the characteristics of text structure are also dissimilar between the two text genres. Narrative texts describe various events, such as a character's goal, action, outcome, and states. Researchers have attempted to identify the structural organizations of narrative texts. Among these theories, *story grammar* (Mandler, 1984; Thorndyke, 1977) and *causal network analysis* (Trabasso & van den Broek, 1985) have been widely used so far.

A story grammar focuses on the typical order of each event that arises in a story. According to Mandler's (1984) story grammar, narrative stories consist of *setting* and *episode*. Stories usually begin with a setting, which introduces characters, time, and locations in the story. The setting is followed by one or more episodes which can be categorized into *beginning*, *development*, *reaction*, *goal*, *outcome*, and *ending*. An episode beginning of an event and it is followed by a development where reaction (e.g., emotional reactions of the character) caused the character to set up a goal, and the character attempt to take actions in order to achieve the goal. These actions lead to outcomes (e.g., success or failure of the attempts) and the story elicits an ending.

Previous studies have examined whether learners are sensitive to story structure in narrative comprehension by adopting story grammar. Mandler (1984) demonstrated that L1 readers' recall reflected the typical order of a story event, based on the story grammar. Moreover, Mandler and Johnson (1977) showed that L1 readers recalled beginnings, goals, and outcomes more than reactions, endings, and actions, suggesting that categories based on the story grammar affect the memory of narrative stories. Also in L2 reading, Horiba, van den Broek, and Fletcher (1993) demonstrated that readers recalled goals and outcomes more than actions and reactions. Thus, information related to goals and outcomes are important elements in narrative comprehension and learning.

Although story grammar reflects the mental structure of narratives constructed by readers, it cannot fully describe and explain how each event causally relates in mental representations. Specifically, a character's action is motivated by the goal, and the story outcome depends on whether this goal is achieved or not (Trabasso & Wiley, 2005). To overcome this shortcoming, some empirical studies examined whether readers identified and represented causal relations among sentences in narrative passages by using causal network analysis (Trabasso & van den Broek, 1985; Trabasso, van den Broek, & Suh, 1989). In this analysis, causal chains of events, actions, and states in a narrative story are described as a network representation. For example, previous studies demonstrated that the events that had more causal connections with other events were recalled better and judged more important than causal "dead ends" (e.g., Trabasso & Sperry, 1985; Trabasso & van den Broek, 1985). Similar findings were observed in research on L2 reading (e.g., Horiba et al., 1993; Ushiro, Shimizu, Kai, Nakagawa, Takaki, Kobayashi, Satake, & Takano, 2010). These results suggested that L2 readers' mental representation of narrative text was influenced by the causal network of a text.

In contrast, the text structure of expository texts varies more than that of narrative texts (e.g., Meyer, Brandt, & Bluth, 1980; Meyer & Freedle, 1984). For example, Meyer and

colleagues categorized expository texts as follows: *collection*, *description*, *causation*, *problem/solution*, and *comparison*. Collection texts list concepts and ideas by association. Description texts provide detailed explanation about a particular attribute, specification, or setting. Causation texts explain causal relationships, such as if-then statements or cause-effect statements. Problem/solution texts present a specific problem, and then present potential solutions to the problem. Finally, comparison texts compare ideas to one another based on the similarities and differences between them.

Some empirical studies have shown that the structure of expository texts affect text comprehension differently. In Meyer and Freedle (1984), L1 students listened to either of four passages with different text structures (comparison, causation, problem/solution, and collection/description) and performed immediate and delayed written recall tasks. The results showed that the comparison and causation structure facilitated learning and memory of the text because these types of text structure are well-organized and provide more cues to retrieve stored information. In L2 reading, Carrell (1984) also examined the effects of expository text structure on reading comprehension and showed that the texts with more organized structure (e.g., comparison, causation, problem/solution) were recalled better than those with less organized structure (e.g., collection of descriptions). In addition, the results of this study revealed that most L2 learners recalled the text with a different structure from the original text. However, Carrell (1992) demonstrated that learners who recalled the text with the same structure as the reading passages showed better performance both quantitatively and qualitatively in the recall task. These results suggest that awareness of the global structure of expository text contributes to better text comprehension.

Thus, narrative and expository texts differ in some aspects. Therefore, to examine the effects of text types on reading comprehension, previous research mainly adopted the following three types of methodology. First, some previous studies compared text memory without

controlling the text content across text types (e.g., Kintsch & Young, 1984; Wolfe, 2005). In Kintsch and Young (1984), L1 readers read and recalled three types of texts (narratives, expository-descriptive texts, and expository-interference texts), the content of which were not controlled; rather, both texts included the same target statements. The results showed that while overall recall performance was best with the narratives, the target recall was better in expository texts than narratives. Wolfe (2005) also demonstrated that narrative texts were better recalled than expository texts when the content was not controlled across two text types.

Second, other studies controlled the content across text types (e.g., Wolfe & Mienko, 2007; Wolfe & Woodwyk, 2010). Wolfe and Mienko presented some factual content in either narrative or expository texts in order to investigate the impact of text genre on learning. The results demonstrated that learning and recall performance did not differ between text types, but the readers with higher prior knowledge benefited more from the expository text compared with the narrative text. Wolfe and Woodwyk also demonstrated that text types influenced not only memory but also on-line processing of to-be-learned content. Specifically, readers generated more prior knowledge elaborations while reading the expository text than the narrative text. These results suggest that readers' prior knowledge was related more to expository comprehension than narrative comprehension.

Other studies did not control the text content itself but controlled reading goals, such as reading news versus literary texts (Zwaan, 1994) or reading for study versus entertainment (Linderholm & van den Broek, 2002; Narvaez, van den Broek, & Ruiz, 1999) in order to examine whether readers altered their processing according to text types. In Zwaan (1994), readers were instructed to read the same text either as a news story or as a work of literature. The results demonstrated that when they read the text as a news story rather than a work of literature, memory of knowledge-based inferences was stronger.

In L2 reading, Horiba (2000, Experiment 1) demonstrated that readers process narrative

and expository texts differently, but differences in processing patterns are rather limited. These results suggested that L2 readers are likely to utilize many cognitive resources in lower-level linguistic processing. Therefore, they engage in language competence-based processing as well as text type appropriate processing. Although Yoshida (2012) also showed the effects of text type on text comprehension, the effects did not appear in total recall production; rather, effects were only found for main idea comprehension in the immediate recall task. These results indicated that the effects of text type on processing and text comprehension were relatively small in L2 reading.

2.2 Establishing Text Coherence Through Inference Generation

2.2.1 Local and Global Coherence of Text Comprehension

Texts are not a simple list or set of isolated words and sentences, rather they are visual communication for transmitting the author's intended messages (e.g., Koda, 2005; Nuttall, 2005). Therefore, writers intentionally write the text in order to share their ideas or messages to readers, and readers attempt to uncover the writers' intended meaning. As reviewed in section 2.1.1, the goal of reading comprehension is to construct coherent mental representations of texts. Therefore, readers need to integrate current information and earlier information, or their background knowledge, in order to build coherence both at local and global levels.

According to Graesser, Singer, and Trabasso (1994), *local coherence* is constructed when "conceptual connections relate the content of adjacent text constituents (i.e., a phrase, proposition, or clause) or a short sequence of constituents (p. 378)." In other words, building local coherence of texts requires readers to understand causal connections between currently processing elements and immediately preceding elements that are represented in short-term working memory. On the other hand, *global coherence* is achieved when "most or all of the constituents can be linked together by one or more overarching themes (p. 378)." Building

global coherence of texts requires readers to link currently processing elements and much earlier elements that are no longer kept in (short-term) working memory and to relevant world knowledge.

Thus, building coherence of texts requires readers to connect and integrate each element in the text into a meaningful representation as a whole. On one hand, readers construct text coherence by using explicit signals, such as connectives (e.g., *before*/*after*, *and*/*but*, *because*, *however*) and co-reference (Halliday & Hasan, 1976). On the other hand, writers do not always explicitly mention the relationships between elements in the text. In such a case, readers need to fill the gaps by inference generation. The term *inference* is defined as "information that is activated during reading yet not explicitly stated in the text" (van den Broek, 1994, p. 556) or "textbase arguments and propositions that were not explicitly mentioned in a message" (Singer, 1994, p. 480). Inference generation is one of the subordinate skills of reading and is regarded as of equal importance to other skills, such as vocabulary, syntax, and semantics (e.g., Grabe, 2009; Grabe & Stoller, 2002; Nuttall, 2005).

2.2.2 Taxonomy of Inferences

The taxonomy of inferences has varied according to the researchers (e.g., van Dijk & Kintsch, 1983; Graesser et al., 1994; McKoon & Ratcliff, 1992; Singer, Graesser, & Trabasso, 1994; van den Broek, Fletcher, & Risden, 1993). Among them, Graesser et al. (1994) categorized inferences into 13 types according to the contents of inference. Table 2.1 shows the classification with a brief description as proposed by Graesser et al. [1]

[1] Van Dijk and Kintsch (1983) proposed *bridging inferences* and *elaborative inferences* in terms of necessity of comprehension. On the other hand, van den Broek et al. (1993) categorized inferences based on the direction of inferences as follows: *connecting inferences, reinstatements, backward elaborations, forward elaborations, orthogonal elaborations*, and *associative inferences*.

Table 2.1

Taxonomy of Inferences (Adapted From Graesser et al., 1994, p. 375)

Type of inference	Brief description
(1) Referential	A word or phrase is referentially tied to a previous element or constituent in the text (explicit or inferred).
(2) Case structure role assignment	An explicit noun phrase is assigned to a particular case structure role, e.g., agent, recipient, object, location, time.
(3) Causal antecedent	The inference is on a causal chain (bridge) between the current explicit action, event, or state and the previous passage context.
(4) Superordinate goal	The inference is a goal that motivates an agent's intentional action.
(5) Thematic	This is a main point or moral of the text.
(6) Character emotional reaction	The inference is an emotion experienced by a character, caused by or in response to an event or action.
(7) Causal consequence	The inference is on a forecasted causal chain, including physical events and new plans or agents. These inferences do not include the character emotions in class (6).
(8) Instantiation of noun category	The inference is a subcategory or a particular exemplar that instantiates an explicit noun or an implicit case role that is required by the verb.
(9) Instrument	The inference is an object, part of the body, or resource used when an agent executes an intentional action.
(10) Subordinate goal-action	The inference is a goal, plan, or action that specifies how an agents' action is achieved.
(11) State	The inference is an ongoing state, from the time frame of the text that is not causally related to the story plot. The states include an agent's traits, knowledge, and beliefs; the properties of objects and concepts; and the spatial location of entities.
(12) Emotion of reader	The inference is the emotion that the reader experiences when reading a text.
(13) Author's intent	The inference is the author's attitude or motive in writing.

According to Graesser et al., classes 1, 2, and 3 are *local inferences* that contribute to building local coherence of texts (between adjacent units), whereas classes 4, 5, and 6 are *global inferences* that contribute to building global coherence of texts. Classes 7 through 11 are *elaborative inferences* that are not necessary for building coherence but lead to deeper understanding of texts. Classes 12 and 13 address the pragmatic communicative exchange between reader and author.

Theories of text comprehension have attempted to explain which inferences are made and under what conditions they are made (e.g., Graesser et al., 1994; McKoon & Ratcliff, 1992; Singer et al., 1994). Inference generation during text comprehension can be accounted for by two reading theories: *the minimalist hypothesis* (McKoon & Ratcliff, 1992) and *the constructionist theory* (Graesser et al., 1994). These theories differ in which types of inferences are generated during reading. Table 2.2 summarized the predictions of on-line inference processing by these two theories.

Table 2.2

Predictions of On-Line Inference Processing by Two Reading Theories

Type of inference	Minimalist hypothesis (McKoon & Ratcliff, 1992)	Constructionist theory (Graesser et al., 1994)
1. Referential	X	X
2. Case structure role assignment	X	X
3. Causal antecedent	X	X
4. Superordinate goal		X
5. Thematic		X
6. Character emotional reaction		X
7. Causal consequence		
8. Instantiation of noun category		
9. Instrument		
10. Subordinate goal-action		
11. State		
12. Reader emotion		
13. Author intent		

Note. This prediction is quoted from Graesser et al. (1994, p. 384). X = on-line prediction.

According to the minimalist hypothesis (McKoon & Ratcliff, 1992), readers do not automatically generate inferences to fully represent the situation described by a text. In the absence of specific or goal-directed strategic processes, inferences of only two kinds are constructed: (a) those that establish locally coherent representations, and (b) those that are rely on information that is quickly and easily available. Therefore, they make the claim that only a few types of inferences (i.e., referential, case structure role assignment, and causal antecedent)

are automatically generated during reading.

On the other hand, the constructionist theory points out that the reader attempts to construct a meaningful situation model that is coherent at both local and global levels. Therefore, according to this theory, inferences which contributed to local and global coherence of the text can be generated during the course of reading comprehension. As such, they claimed that six types of inferences (i.e., referential, case structure role assignment, causal antecedent, superordinate goal, thematic, and character emotional reaction) are on-line inferences. However, Graesser et al. (1994) claimed that readers do not construct globally coherent representations of the text when the following conditions are met: (a) if the reader regards the text as "inconsiderate" (i.e., lacks global coherence and a message), (b) if the reader does not have enough background knowledge, and (c) if the reader has goals that do not require the construction of a meaningful situation model (e.g., proofreading for spelling errors). In other words, when there is a breakdown in these conditions, the readers construct only local coherence of texts or do not try to achieve any coherence at all.

Thus, the minimalist hypothesis and constructionist theory are in disagreement, especially in generation of inferences that establish global coherence. Specifically, while the minimalist hypothesis claims that global inferences are only generated strategically and not automatically, the constructionist theory claims that global inferences are generated spontaneously. However, while research on local inferences that are elicited by one or two sentences thus far have received more research attention, less attention has been paid to generation of global inferences. Therefore, the current research examined the construction of global coherence of texts, focusing on *thematic inference* in narrative reading and *superordinate inference* in expository reading. The following sections review the establishment of coherent mental representations of narrative texts (section 2.2.3) and that of expository texts (section 2.2.4). In addition, some previous L2 reading studies regarding inference generation are

reviewed.

2.2.3 Building Coherence of Narrative Texts Through Thematic Inference

When people read a text, they derive meaning from it by comprehending individual words and sentences. A narrative text usually contains diverse information such as beginning, setting, action, goal, attempt, outcome, and ending (e.g., Thorndyke, 1977). However, readers need to comprehend not only explicit information in the text but also messages conveyed by the writer to comprehend the text deeply. The overall meaning, message, or point of a text is called the *theme* (e.g., Graesser, Pomeroy, & Craig, 2002; Graesser et al., 1994; Kurz & Schober, 2001).

According to Graesser et al. (2002), the term *theme* has been mainly used from two different aspects: *theme-topic* and *theme-motif*. The theme-topic is a content word about the topic of the text and refers to the subject matter of a passage. For example, the title of the text could be a theme-topic. On the other hand, the theme-motif is "a point, message, or moral expressed as a declarative statement" that is expressed as an adage. For example, the theme-topic of the following passage could be "proposal of marriage," whereas the theme-motif could be its moral, which is "it is too late to try to prevent something after you have noticed."

> Phil was in love with his secretary and was well aware that she wanted to marry him. However, Phil was afraid of responsibility, so he kept dating others and made up excuses to postpone the wedding. Finally, his secretary got fed up, began dating, and fell in love with an accountant. When Phil found out, he went to her and proposed marriage, showing her the ring he had bought. But by that time, his secretary was already planning her honeymoon with the accountant. (Adopted from Seifert et al., 1986; p. 231)

In that thematic inference is the main interest of the current research, the following

discussion focuses on theme-motif rather than theme-topic, because thematic inference is defined by Graesser et al. (1994) as follows: "integrate major chunks of the text or that convey the point of a message. For example, a story might be an instantiation of the virtue 'practice what you preach'" (p. 372).

Some previous L1 research has empirically investigated whether thematic inferences are generated during reading. For example, some studies have employed priming tasks in which the participants decided, as quickly as possible, whether the test item was an appropriate thematic inference word generated by the text (Long, Oppy, & Seely, 1994, 1997; Till, Mross, & Kintsch, 1988). They concluded that thematic inference was not activated automatically during reading. However, Zhang and Hoosain (2005) pointed out that these studies ignored the reader's processing time for individual words in the text; that is, if the learners had been given enough time when presented with the text and target words in the priming task, they would have automatically activated thematic inferences. Finally, they demonstrated automatic activation of thematic inferences by controlling the task condition (Zhang & Hoosain, 2005) and using a self-paced reading time method (Zhang & Hoosain, 2001).

Although some of these studies supported the notion of constructionist theory, in that thematic inferences were activated automatically during reading, there was a major methodological problem. For example, Zhang and Hoosain (2005), which targeted native speakers of Chinese, presented single-character Chinese words (e.g., greedy, succeed, kindness) as target words in the lexical decision tasks. Presenting single words can be considered as an inappropriate methodology because of the following two points. First, regarding the original definition of thematic inference (Graesser et al., 1994), single words are rarely adequate for narrative themes because they do not fully convey the writer's message or points. Second, single words might be easily activated through associative processes (see the construction-integration model and the landscape model in section 2.1.1) rather than connecting global units of texts;

therefore, the single-word priming or lexical decision tasks do not directly reflect thematic inference generation during reading. Zwaan, Radvansky, and Whitten (2002) proposed a similar claim as follows: Single words (that express theme-topics) can be generated without the prior construction of a situation model, and they can be inferred based on textual markers rather than a situation model. Therefore, it can be difficult to conclude generation of global inference, including thematic inference, from the studies that adopted single-word priming or the lexical decision task.

In another study, Seifert et al. (1986) also used a priming procedure but examine memory connections between story pairs. Two types of story pairs were prepared: (a) same-theme condition in which the pair of stories shared the same themes, and (b) different-theme condition in which the stories were based on different themes. As priming sentences, they presented the conclusion of test sentences and the reaction times between the two reading conditions were compared. They found facilitation for targets in the same-theme condition to the different-theme condition. However, this effect was found only when the participants were asked to attend the thematic similarity of the stories (i.e., reading for judging thematic similarity between texts). Therefore, they concluded that thematic inference was generated through strategic processing which is induced by task instructions, rather than automatically activated during reading.

Kurtz and Schober (2001) more directly examined thematic inference during reading by requiring readers to think-aloud their thoughts at specific points of reading. The results indicated that thematic inferences are not activated at the moment of initial comprehension, rather they are constructed later as acts of interpretation. Therefore, they also concluded that thematic inference is not on-line automatic processing, rather strategic processing occurred in off-line products.

Some empirical studies demonstrated possible factors that affect thematic inference generation (e.g., Dorfman & Brewer, 1994; Lehr, 1988; Narvaez, 1998, 1999; Whitney, Ritchie,

& Clark, 1991; Zhang & Hoosain, 2001). Narvaez (1998) examined whether third- and fifth-grade children and university students can understand a theme through reading a story. The participants were told to read a story aloud to answer the message of the story. After reading, they were required to rate how a set of vignettes matched the story theme using a 5-point Likert scale, and to select the best theme for the story by multiple-choice questions. The results demonstrated that younger children were likely to select the inappropriate theme. Moreover, story comprehension measured by true-false questions significantly related to theme comprehension. In addition to off-line theme comprehension, Narvaez (1999) examined the processes of reading moral stories by adopting the think-aloud method. The study compared expert and less-expert readers, and demonstrated that the expert readers indicated deeper engagement in the texts (e.g., more explanations, predictions, and evaluations) than less-expert readers. These results suggested that theme comprehension through reading stories were influenced by readers' characteristics.

Moreover, Whitney et al. (1991) demonstrated that readers with a high capacity of working memory can maintain local coherence while activating possible thematic inferences. On the other hand, readers with a low capacity of working memory tend to focus on a sentence-to-sentence understanding and activate specific thematic inferences at an early stage in reading; therefore, it is difficult for them to change their representation of the whole text during the reading process.

The other factor that appears to affect thematic inference generation is contextual information in the text. Some previous studies have regarded a *central action* and its *outcome* in narrative texts as important components for theme comprehension (Dorfman & Brewer, 1994; Zhang & Hoosain, 2001). Dorfman and Brewer (1994) demonstrated that readers have difficulty in understanding the theme of a text when the protagonist's central action and outcome are not consistent (e.g., when a negative action causes a positive outcome). Zhang and

Hoosain (2001) observed the two text factors affected thematic inference generation. Experiment 1 in their study demonstrated that presenting the title of narrative texts facilitated thematic inference generation, suggesting that the title helped readers to construct globally coherent representations. In Experiment 2, they examined the effects of context on thematic inference generation by manipulating the consistency of the protagonist's goal and its outcome. The results indicated that readers generated thematic inferences when the goal and its outcome were consistent.

In sum, thematic inference, which is defined as a point, message, or moral of the text, is constructed through reading narrative texts by L1 readers; however, its activation can be through strategic processing rather than automatic and spontaneous. Moreover, both reader and text factors can influence thematic inference generation.

2.2.4 Building Coherence of Expository Texts Through Superordinate Inference

Compared to previous studies on narrative comprehension, studies on inferences during reading of expository texts are few (Lorch, 2015 for a review). Among the studies on inference generation in expository texts, many studies focused on causal inferences during reading (e.g., Noordman, Vonk, & Kempff, 1992; Singer, Harkness, & Stewart, 1997). For example, Singer et al. (1997) demonstrated that L1 readers made causal inferences when they read simple scientific texts with familiar topics, whereas such inferences were not generated when reading difficult texts. A similar finding was observed in Noordman et al. (1992) when readers failed to generate causal inferences in expository reading unless they were given strategic instructions that encourage inference generation. Furthermore, Wiley and Myers (2003) also demonstrated that readers did not generate causal inferences if a single filler sentence was inserted between target sentences. This result suggested that even L1 readers had difficulty maintaining global coherence of expository texts. According to Lorch (2015), making causal inferences in

expository reading was more difficult than in narrative reading due to less prior knowledge and the greater complexity of expository texts (as reviewed in section 2.1.2), which leads to difficulty in strategic processing and metacomprehension of texts.

Lorch and van den Broek (1997) claimed that many studies have focused on how readers construct horizontal connections between events in the text (i.e., one-to-one local relations) and not on vertical relations, which contribute to global coherence of text representations. Although few studies have directly focused on the construction of globally coherent representations of expository texts, some previous research provides information relevant to the current research. For instance, *superordinate inference*s are inferences about how several ideas or events are connected in the text (i.e., concept-to-concept relations), such as the relation between a sequence of statements and a script or scheme that subsumes them, and between a sequence of statements and the generalization that subsumes them (e.g., Lorch & van den Broek, 1997; Ritchey, 2011).

The process of drawing superordinate inference can be explained by Kintsch and van Dijk's theory. In text comprehension, readers first decode each word to construct meaning, and finally organize them into a higher-level representation which is the gist of the text. Kintsch and van Dijk (1978) called these two components of text comprehension the *microstructure* and the *macrostructure*. The microstructure is the local structure of the text, the sentence-by-sentence information, as supplemented by and integrated with long-term memory information. The macrostructure is a hierarchical network of propositions that captures the main idea, or theme, of a discourse; therefore, to construct coherent representation of the overall text, constructing macropropositions as well as micropropositions are required. According to Kintsch (1998), texts are composed of textbase information which is a network of propositions, and each piece of information has hierarchical relationships with other information. Therefore, readers form an overall text representation through the comprehension of the relationship

between textual information. Figure 2.1 illustrates a hierarchical structure of micropropositions and macropropositions. As the reader proceeds through a text, he or she builds the microstructure, step by step, by applying relations of local coherence (e.g., referential, causal, and temporal). In the next step, a more global kind of coherence in discourse (i.e., the macrostrucrure) is built by reorganizing the microstructure into a coherence global structure in terms of meaningful units that account for the gist of the text (Tapiero, 2007).

In the process of the construction of macropropositions, readers recognize the importance and the hierarchical level of each element in the text. Based on this process, readers select important elements in the text and integrate them to generate superordinate propositions. This process of generating superordinate propositions includes the following three rules called *macrorules* (Kintsch & van Dijk, 1978): (a) deletion (the omission of a proposition that is unimportant and of irrelevant information), (b) generalization (the conflation of details into higher-level categories), and (c) construction (the integration of details into topic sentences). Thus, through this process, readers form the "gist" that represents their overall comprehension of the text.

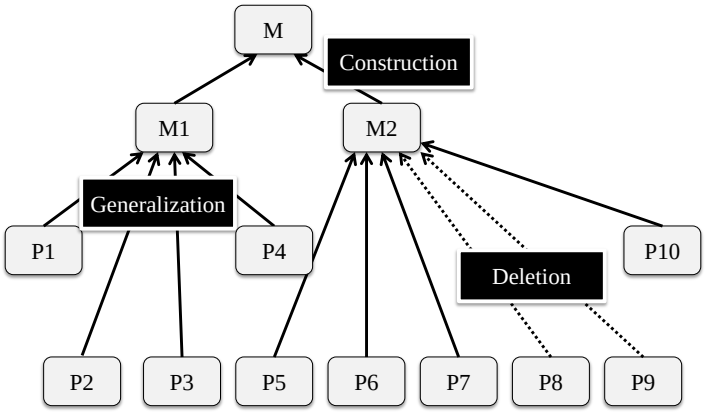

Figure 2.1. The macroprocessing during reading. P = microproposition; M = macroproposition.

When the superordinate macroproposition is explicitly stated as a title or topic sentence of the text, it plays a role of an organizational cue for the reader. As a result, readers use it to understand the gist of a text, which leads to better text comprehension (e.g., Lorch, Lorch, & Inman, 1993; Sanchez, Lorch, & Lorch, 2001). In Lorch et al. (1993), L1 readers read a text with or without signaling devices (e.g., headings, overviews, and summaries) and recalled the topics and content of the text they had read. The results demonstrated that signaling devices facilitated readers in remembering more topics. Furthermore, readers' recall was more organized when they read the text with signals than without. Sanchez et al. investigated the effects of headings on readers' text memory and the trainings for the structure strategy. The results showed that if the participants either received training in the structure strategy or read a text with headings, memory of the text was better than if they did not receive training or headings.

However, the macroproposition that subsumes a sequence of statements is not always explicitly stated in the passage. In that case, readers need to generate superordinate inference in order to construct the macroproposition. Some empirical studies examined whether readers inferred the implicit macropropositions that connects statements across an expository text (e.g., Brown & Day, 1983; Brown, Day, & Jones, 1983; Ritchey, 2011; Williams, Taylor & de Cani, 1984). Although in these studies, the term "superordinate inference" was not used and some researchers did not directly focus on inference generation, they provide important findings regarding the construction of globally coherent representation of expository texts.

For example, Williams et al. (1984) asked L1 readers to write a summary sentence of short expository paragraphs consisting of three sentences (e.g., *Cowboys had to protect the herd from cattle robbers. They had to brand cattle to show who owned them. Sometimes cowboys had to separate the cattle that were to be sent to market*.). In this study, a summary sentence, such as "Cowboys had jobs to do," was regarded as an appropriate macroproposition, which

subsumes all three sentences. Results demonstrated that immature readers (third, fifth, and seventh graders) tend to write a sentence that was too narrow in focus, suggesting difficulty with integrating each sentence to construct the overall image of the text.

Similar findings were also found when assessing comprehension of longer texts. In Brown et al. (1983), L1 readers (fifth, seventh, and eleventh grade, and college students) completed summary tasks about the texts they had read. The researchers found that students used a "mature strategy" where they combined across paragraphs in order to express the gist of large bodies of text. However, such a strategy was observed only in college students. Brown and Day (1983) also examined the relationships between readers' development and summary ability. It was demonstrated that the ability to construct implicit macroproposition developed with age. These results suggest that understanding implicit macropropositions by integrating information across paragraphs was especially difficult for less skilled readers.

In a recent study, Ritchey (2011) directly examined generation of superordinate inferences. In this study, college students read expository texts that contained the following target statements: (a) *consistent*, which described consistency with both the generalization implied by the text and the actual topic of the text, (b) *inconsistent*, which described inconsistency with the generalization implied by the text but consistency with the actual main topic, and (c) *off-topic*, which described inconsistency with both the generalization implied by the text and the topic of the text. The readers were assigned to read half of the text with the summarization task, which aimed to encourage processing relations among sentences, and the other half of the text with the verification task, which aimed to encourage processing details. This study assumed two possibilities as follows. If superordinate inference was an option and strategic processing occurred under specific conditions, the task instructions would affect the reading time on target statements. On the other hand, if superordinate inference was a routine and automatic processing, the task instructions would not affect the reading time. The results

showed reading times were shorter on consistent statements than inconsistent statements, and on inconsistent statements than off-topic statements. Moreover, such reading time patterns were observed regardless of the task instructions, suggesting that superordinate inferences were a mandatory part of expository reading comprehension among L1 readers. Effects of task instructions are further reviewed in section 2.3.

2.2.5 Inference Generation in L2 Reading Comprehension

Although some researchers have investigated inference generation among L2/EFL readers (e.g., Horiba, 1996, 2000; Muramoto, 2000; Nahatame, 2014; Shimizu, 2015; Yoshida, 2003), most of these studies examined inference generation in narrative texts.

Some L2 reading research adopted a *think-aloud method* in which the readers were asked to talk about what they thought was happening, in order to examine inference generation while reading narrative texts (e.g., Horiba, 1996; Yoshida, 2003). Horiba (1996) compared L1 and L2 readers' processes and comprehension of narrative passages. The results showed that L2 readers allocated more cognitive resources to lower-level processing (e.g., graphomorphemic/graphophonemic analysis, word recognition, and syntactic/semantic analysis) compared to L1 readers. Furthermore, while L2-advanced readers made both backward and forward inferences during reading, L2-intermediate readers did not. Yoshida (2003) also employed the think-aloud method and found a difference between proficient and less-proficient L2 learners in inference generation. She reported that high-proficiency readers generated more elaborative inferences during reading than lower-proficiency readers. Additionally, lower-proficiency readers had difficulty in elaborative inference generation because of the inadequateness of their lower-level processing.

Muramoto (2000) investigated what types of inferences (i.e., goal, action, emotion, and state) are generated and how learners' L2 proficiency had an influence. He adopted a *sentence*

recognition task where participants were required to judge whether the target sentences had appeared in the text they had just read. If they made inferences and encoded it into their text representation, they should falsely recognize that the target sentences, which were not explicitly stated in the text but can be inferred from the text, were written in the text. The results demonstrated that the recognition rate for inference statements was higher in the proficient learners than in the less proficient learners, suggesting that proficient learners generated more inferences than less proficient learners, as demonstrated in other L2 reading research (e.g., Horiba, 1996; Yoshida, 2003). Moreover, the recognition rate was higher in goal and action inference statements than emotion and state inference statements. He claimed that such differences in recognition rate by inference types can be attributed to the differences in the characteristics of inferences. Goal and action inference play roles in connecting information in passages; therefore, the necessity of these inferences for text comprehension was high. On the other hand, while emotion and state inferences embellish the representations of narrative texts they are not necessary for text comprehension.

Other L2 studies demonstrated that inference generation was influenced not only by factors pertaining to readers, but also textual factors (Barry & Lazarte, 1998; Horiba, 1996). Barry and Lazarte examined the effects of domain-related knowledge, syntactic complexity of texts, and the topic of texts on L2 readers' inference generation (within-text inference, elaborative inference, and incorrect inference). To test the effects of syntactic complexity, they manipulated the number of embedded clauses per sentence included in the experimental passage. The results of a written recall task showed that when the syntactic complexity increased, L2 learners with less prior knowledge had difficulty maintaining prior elements in the text, which lead to minimal and inaccurate inference generation. Horiba (1996) also investigated the effects textual coherence (low- and high- coherence texts) on inference generation by manipulating the number of causal links between events in narrative texts. One of the interesting findings of this

study was that while L1 readers generated more elaborations for low-coherence texts than for high-coherence texts, L2 readers did not change their processes according to the coherence. It was discussed that the reason for this might be L2 learners' lower standards of coherence than L1 readers.

While the aforementioned studies examined several types of inferences during reading, other studies have focused on more specific types of inference. For example, Ushiro et al. (2012) investigated the activation and encoding of two types of inferences, bridging and predictive inferences, by using an on-line probe recognition task and an off-line recall task. The results demonstrated that bridging inferences were activated on-line, whereas predictive inferences were generated with some delay. These results were consistent with the minimalist hypothesis (McKoon & Ratcliff, 1992) and constructionist theory (Graesser et al., 1994). Nahatame (2014) narrowed his focus on generation of predictive inferences and confirmed that L2 readers did not generate predictive inferences during reading when they were instructed simply to read the passages for comprehension. Moreover, it was demonstrated that L2 readers' predictive inference generation was facilitated by strategy instructions. The effects of tasks and strategy instructions on reading processes and text comprehension are further reviewed in the next section.

However, the main research interest of the aforementioned studies was to examine how inference generation contributes to building local coherence of texts rather than global coherence. Shimizu (2015) focused on Japanese EFL learners' generation of local bridging inferences and global bridging inferences during reading of narrative and expository texts. The main findings regarding their research were as follows. First, the results of think-aloud data demonstrated that lower-proficiency readers tended to allocate more cognitive resources to lower-level processing in expository reading than in narrative reading. Second, while local bridging inferences co-occurred with paraphrasing and rereading the previous sentences as in

L1 previous studies, global bridging inferences co-occurred with several processes.

Although the following studies did not directly focus on inference generation in L2 reading, they provide findings regarding the construction of global coherence of texts. For example, Ushiro, Nakagawa, Kai, Watanabe, and Shimizu (2008) used an expository text with four paragraphs and manipulated the explicitness of macroproposition in the text, and college students completed a summary task for the text. The results indicated that while the learners had the ability to construct macropropositions of a single paragraph, they lacked the ability to construct a macroproposition that embraces ideas of more than one paragraph. Moreover, although the learners identified the hierarchical structure of the expository text regardless of the macropropositions, the elimination of macropropositions hindered the connection of some of the information in the text.

Morishima (2013) aimed to examine whether L2 readers maintained global coherence of narrative texts by using the *inconsistency detection paradigm*. The main findings of this study were that while L2 readers maintained local coherence of the texts as L1 readers did, they had difficulty constructing global coherence of the texts due to the limited resource allocation for discourse-level processes.

Because a very small number of studies have directly examined the construction of globally coherent representations of texts in L2/EFL reading comprehension, it is not possible to derive clear predictions for the current research. However, with regard to the results of previous literature, the following three points can be summarized. First, it can be difficult, even impossible, for Japanese EFL learners to generate thematic inferences in narrative texts and superordinate inferences in expository texts due to limited linguistic knowledge and reading skills. Second, as with other inference types, thematic and superordinate inferences can be influenced by learners' L2 proficiency. Third, given the complexity and characteristics of expository reading, generating superordinate inferences in expository texts can be more

difficult than generating thematic inferences in narrative texts for EFL learners.

Thus, inference generation for establishing global coherence of texts can be difficult for Japanese EFL learners and can be influenced by reader and text factors. In addition, the extent to which readers engage in processes to construct a coherent representation depends on the reading goal given by the task instructions. In other words, the readers' standards of coherence are affected by the task instructions (e.g., McCrudden, Magliano, & Schraw, 2010; van den Broek, Lorch et al., 2001). Therefore, the next section reviews effects of task instructions on reading processes and text comprehension in L1 and L2 reading.

2.3 Effects of Task Instructions on Reading Comprehension

People read texts for a variety of reasons, such as for entertainment, studying, or recall; therefore, reading comprehension is conceptualized as a goal-directed activity in which the reader uses text to accomplish some task (McCrudden et al., 2010). How reading goals affect inference generation in reading comprehension can be accounted for by the concept of standards of coherence (van den Broek et al., 2001; van den Broek et al., 1995). As readers process a text, they have their own standards that act as criteria for comprehension. Readers systematically alter their criteria for comprehension according to different reading situations (e.g., the text genre, reading task, motivation) and generate different patterns of inference. Thus, the task instructions determine what types of coherence (e.g., referential, causal, spatial, temporal, and logical) should be maintained in comprehension, and what types of inferences are needed for it.

2.3.1 Effects of Task Instructions on L1 Reading Comprehension

Many previous L1 studies have revealed that the task instructions given before reading affect reading processes and post-reading text comprehension (e.g., Bråten & Strømsø, 2009; Bohn-Gettler & Kendeou, 2014; Linderholm & van den Broek, 2002; Magliano, Trabasso, &

Graesser, 1999; McCarthy & Goldman, 2015; Narvaez et al., 1999; van den Broek, Lorch et al., 2001). Table 2.3 shows the summary of previous studies which investigated effects of task instructions on processes and products of L1 reading comprehension.

For example, some studies demonstrated that L1 readers used different cognitive processes and strategies when reading for study versus when reading for entertainment (e.g., Bohn-Gettler & Kendeou, 2014; Linderholm & van den Broek, 2002; Narvaez et al., 1999; van den Broek et al., 2001). Specifically, when reading for study, readers used strategies such as generating inferences, paraphrasing, and rehearsal of text information; and these processes resulted in better text comprehension than reading for entertainment. Conversely, when reading for entertainment, readers generated associations and formed opinions or gave comments about the text content.

Magliano et al. (1999) examined the effects of more specific types of strategic instructions (i.e., read passages for explanation, prediction, association, or understanding) on readers' strategy use and text comprehension. The analysis of think-aloud protocols indicated that L1 readers were able to strategically control their inference generation according to the given instructions. Specifically, compared to the readers in the understanding condition, the readers in the explain condition produced more comments on explanations, those in the prediction condition generated more predictive inferences, and those in the association condition produced more associations.

Table 2.3

Summary of Previous Studies Investigating Effects of Task Instructions on Processes and Products of L1 Reading Comprehension

Study	Participants	Texts	Instructions	Methodology	Effects on Processes	Effects on Products
Bråten & Strømsø (2009)	184 Norwegians students, mean age of 22.6	seven separate texts about climate change	Argument, Summary, Global understanding	Inference verification task	N.A	Readers in argument and summary conditions performed better in inference verification task than those in global understanding condition.
Bohn-Gettler & Kendeou (2014)	83 native English-speaking undergraduates	16 expository texts (compare-contrast, problem-response, description, and chronological)	Reading for study vs. reading for entertainment	Think-aloud task, Summary recall	The entertainment condition engaged in non-coherence processes (e.g., associations, comments) more than the study condition. Readers with high working memory engaged in more paraphrasing and connecting inferences in the study condition than the entertainment condition.	Readers with the goal of studying performed better in summary recall than those with the goal of entertainment.

Linderholm & van den Broek (2002)	110 native English-speaking college students	Two expository texts	Reading for study vs. reading for entertainment	Think-aloud task Written recall task	The entertainment condition engaged in non-coherence processes (e.g., associations, opinions) more than the study condition. The study condition engaged in more connecting inferences and paraphrasing than the entertainment condition. Readers with high working memory engaged in less demanding processes which were not beneficial for text comprehension.	Readers with low working memory capacity recalled the same amount across reading conditions, whereas readers with high working memory capacity recalled more in the study condition than those in the entertainment condition.
Magliano et al. (1999, Experiment 1)	48 native English-speaking college students	Eight short narrative stories	Explain, Predict, Associate, Understand	Think-aloud task, Written recall task	Readers were able to strategically control the inferences that they generate according to the given instructions.	There are no differences in recall performance as a function of reading goals.
McCarthy & Goldman (2015, Experiments 1 and 2)	114 native English-speaking college students, mean age of 19.9	A science-fiction story with 2,201 words	Plot, Ambiguous, Argument, Theme	Essay writing	N.A	Readers in plot condition generated more paraphrased information in their essay than the other conditions. Readers in argument and theme conditions produced more text-based inferences and interpretive inferences than other conditions.

Narvaez et al. (1999)	20 native English-speaking undergraduates, mean age of 23.1	two narrative literary texts and two expository texts	Reading for study vs. reading for entertainment	Think-aloud task, Reading time, Written recall task, Comprehension questions, Questionnaire on reading strategies	Readers in the study condition engaged in more repeating and evaluating processes than those in the entertainment condition.	There are no differences in recall performance and comprehension questions as a function of reading goals.
van den Broek et al. (2001)	82 native English-speaking undergraduates	four expository texts with problem-solving structure	Reading for study vs. reading for entertainment	Think-aloud task, Written recall task	Readers in the study condition generated more coherence processes (e.g., explanatory inferences, predictive inferences) than those in the entertainment conditions. Readers in the entertainment condition generated more associations than those in the study condition.	Readers with the goal of studying performed better in the recall task than those with the goal of entertainment.

Thus, many studies provide evidence that L1 readers are able to adjust their cognitive processes and strategies in accordance with their reading goals. On the other hand, some studies demonstrated that flexible control of resource allocation can happen only in high-working memory readers or skilled readers (Bohn-Gettler & Kendeou, 2014; Linderholm & van den Broek, 2002). For example, Linderholm and van den Broek demonstrated that readers with high-working memory capacity are better able to control their inferential processing than those with low-working memory capacity when they are instructed to read the passage for study.

2.3.2 Effects of Task Instructions on L2 Reading Comprehension

Within the field of L2 or EFL reading, some researchers have examined the effects of task instructions on text processing and comprehension (e.g., Horiba, 2000, 2002, 2013; Nahatame, 2014; Yoshida, 2012). Table 2.4 shows the summary of previous studies which investigated effects of task instructions on processes and products of L2 reading comprehension.

Horiba (2013) investigated the effects of three types of task instructions: to pay attention to words and expressions in the text (Expression condition); to visualize in their minds events, states, and actions in the text (Image condition); and to compare the author's views with their own views and evaluate them (Critique condition). The results of Experiment 2, which adopted the think-aloud technique, indicated that L2 learners' processes were partly changed by the task instructions. Specifically, the learners in the Expression condition engaged in lower-level linguistic processing than those in the other conditions. The learners in the Critique condition made more comments on higher-level conceptual processing (e.g., reaction and evaluation) than learners in the other conditions. On the other hand, the results of a written recall task demonstrated that the task instructions did not affect text comprehension.

Similar results were found in a study by Yoshida (2012) that compared three types of task instructions: outlining, answering embedded questions, and reading only. The results did not

show the effects of task instructions on both immediate and delayed text comprehension. Moreover, Horiba (2002) also suggested that a reading goal sometimes produces negative effects on text comprehension due to the overload of task demands (e.g., reading for critique) caused by limited language proficiency.

Table 2.4

Summary of Previous Studies Investigating Effects of Task Instructions on Processes and Products of L2 Reading Comprehension

Study	Participants	Texts	Instructions	Methodology	Effects on Processes	Effects on Products
Horiba (2000, Experiment 2)	14 native English-speaking readers and 14 native Japanese-speaking readers	Two short news paper articles	Reading freely vs. Reading for coherence	Think-aloud task, Summary recall task	L1 readers generated more backward inferences, forward inferences and knowledge associations in the read-for-coherence condition than the read-freely condition. In L2 reading, there were no differences in think-aloud protocols between the two reading conditions.	There was no effect of task condition on L1 readers' recall. L2 readers in the read-for-coherence condition performed better in summary recall than those in the read-freely condition.
Horiba (2002)	84 native Japanese-speaking undergraduates, mean age of 19	Two short expository texts (Japanese, English)	Read for surface forms, Read for meaning, Read for critique	Written recall task	N.A	There was no effect of task condition on L1 readers' recall. L2 readers in the read-for-critique condition performed worse in recall task than those in the other conditions.
Horiba (2013, Experiment 2)	28 native Japanese-speaking undergraduates, mean age of 20.3	An argumentative essay in English	Expression, Image, Critique	Think-aloud task, Written recall task	Readers in the Expression condition produced more comments on the structure analysis than the Critique condition. The Critique condition produced more comments on responses (e.g., association, evaluation) than the Expression condition.	There were no significant differences in recall performance between task conditions.

Nahatame (2014)	40 native Japanese-speaking undergraduate and graduate students, mean age of 20.55	32 short narrative texts	Predict, Control	Lexical decision task, Sentence reading times, Written recall task	Providing strategy instructions facilitated predictive inference generation.	Readers in the strategic orienting condition performed better in the recall task than those in the non-orienting condition.
Yoshida (2012)	103 native Japanese-speaking undergraduates	An expository text, a narrative text	Outlining, Answering embedded questions, Reading only	Written recall task (immediate and delayed)	N.A	There were no significant differences in recall performance between task conditions. (A slight qualitative difference was found in the recall performance.)

Thus, the effects of task instructions did not affect or even hinder L2 learners' text comprehension. Horiba (2013) claimed that the learners were given a particular task but they were not told to engage in a specific strategy or use processing behaviors during reading; therefore, clear effects of task instructions were not found in text comprehension. Indeed, some L2 studies proved that specific types of strategic instructions did facilitate text comprehension. Horiba (2000, Experiment 2) compared participants' reading comprehension in a condition where they were assigned task instructions to read for coherence (i.e., think about how information in the current sentence is related to prior text and how it might be related to later text) versus a read freely condition. The results demonstrated that the task instructions facilitated their recall production rates, suggesting that L2 readers constructed more coherent representations when explicitly instructed to do so. Another important finding was that L2 readers controlled their reading processes less flexibly than L1 readers; that is, for L2 readers, changing a reading strategy according to a reading goal or task is more difficult than for L1 readers. Nahatame (2014) also examined the effects of specific strategy instructions to predict likely outcomes in text comprehension. The results revealed that the strategy instructions facilitated learners' predictive inference generation, and such strategic processing also improved text comprehension.

The results of these L2 studies suggest that the effects of reading goals on text comprehension seem to be more complex and unstable for L2 and EFL reading than for L1 reading. On the other hand, the following two possibilities can be derived from the previous studies. First, as in Horiba (2000) and Nahatame (2014), when learners are given task instructions that require them to alter specific types of processing during reading (e.g., task instructions aimed at a specific type of inference), they will be able to strategically control their inference generation according to the given instructions, which leads to a positive influence on text comprehension. Second, some previous studies adopting think-aloud methods proved that

L2 readers were not likely to change their cognitive processes during reading in accordance with reading goals. However, their standards of coherence, in terms of the types of coherence that should be maintained in comprehension, might be changed (e.g., referential, causal, spatial, temporal, and logical).

2.4 Limitations of Previous Studies and Purpose of Current Research

As reviewed in this chapter, the core of reading comprehension is to construct coherent mental representations of texts. Therefore, many previous studies have examined the generation of various types of inference and developed theories regarding reading comprehension. However, there are some limitations and issues to resolve with regard to the previous studies, which can be summarized in the following three points.

First, compared to the inferences that establish local coherence of text, few research studies have focused on the inferences establishing global coherence of text, especially in L2 reading research. The researchers have paid more attention to local inferences than global coherences, because manipulating and controlling experimental material is easier in experiments that target local coherences. However, given that discourse is not a set of isolated words and sentences, rather it conveys the messages intended by writers, building globally coherent representations is important for uncovering the messages through reading the text as a whole. Therefore, global coherences should gather more research attention in order to reveal reading comprehension among L2/EFL learners. Although some L2 reading research investigated inference generation during reading (e.g., Horiba, 1996; Muramoto, 2000; Yoshida, 2003), few studies examined the specific types of inferences, especially global inferences. Therefore, the current research aimed to investigate Japanese EFL learners' construction of globally coherent representations of texts, focusing on thematic inference in narrative comprehension and superordinate inference in expository reading.

Second, although many theories regarding reading comprehension assume the interaction between texts, readers, and tasks on text comprehension, the number of L2 studies examining such interactions is still small (e.g., Horiba, 2000, 2013; Yoshida, 2012). Therefore, the design of the current research assumed that text genre (narrative and expository texts), readers' L2 reading proficiency (upper and lower proficiency levels), and task instructions (read for comprehension, read for constructing global coherence) could affect inference generation and text comprehension. Investigating such interactions among these three factors will provide meaningful implications for educational settings.

Third, only a limited number of studies have discussed reading comprehension from the viewpoint of both the processes and products of comprehension (e.g., Horiba, 2000, 2013; Nahatame, 2014). To clarify the characteristics of EFL readers, not only the products of reading (i.e., what the learners understand from the text after reading) but also the processes of reading should be tested, as well as the relationships between the two.

To resolve these issues, the current research conducted a total of six experiments as follows. First, Study 1 of this dissertation involved three experiments (Experiments 1 to 3) to examine thematic inference generation in narrative reading. Although the focus in some of the previous research was on-line generation of themes during reading, the methodologies used in the studies could not directly show generation of thematic inference. Therefore, Experiment 1 investigated whether Japanese EFL readers are able to generate thematic inference, by directly comparing readers' comprehension of explicit and implicit themes and comparing the themes with other types of inferences. Experiment 2 focused on the effects of L2 reading proficiency and task instructions on thematic inference generation and text comprehension. Experiment 3 focused on reading processes as well as products after reading, considering the interaction between L2 reading proficiency and the task instructions.

Study 2 involved three experiments (Experiments 4 to 6) to investigate superordinate

inference generation in expository reading. The design of the experiments was similar to that of Study 1. Experiment 4 investigated whether Japanese EFL readers are able to generate superordinate inference, by directly comparing readers' comprehension of explicit and implicit superordinate propositions. Experiment 5 focused on the effects of L2 reading proficiency and task instructions on superordinate inference generation and text comprehension. Finally, Experiment 6 explored reading processes as well as products after reading.

As overviewed above, the current research examined global coherence of mental representations in narrative and expository reading respectively, because inference types that contribute to building globally coherent representations differ between text types. However, given the significance of the interaction between texts, readers, and tasks in reading comprehension, this dissertation discussed these interactions by combining the results derived from six experiments.

Chapter 3

Study 1: Thematic Inference Generation in Narrative Reading

3.1 Experiment 1: Understanding Implicit Themes Through Inference Generation in Narrative Reading

3.1.1 Purpose and Research Questions

The purpose of Experiment 1 is to examine whether Japanese EFL learners can understand the implicit theme of narrative texts through inference generation. To investigate readers' inference generation, previous studies directly compared two types of passages in which specific information was explicitly or implicitly stated (e.g., Poynor & Morris, 2003; Ushiro, Nahatame, Hasegawa, Kimura, Hamada, & Tanaka, 2014). Therefore, based on these previous studies, the present study manipulated the explicitness of the theme in experimental passages.

In order to examine thematic inference generation, the present study adopted an *inference verification task*, where participants were required to judge whether target statements can be understood or suggested from the passage they had read, instead of whether the statements appeared in the passage (i.e., sentence recognition).[2] The inference verification task is a valid measure of deeper, situational understanding of texts and has been widely used in previous studies to test the ability to make inferences (Bråten & Strømsø, 2009; Campion & Rossi, 2001; Rapp & Gerrig, 2006; Royer, Carlo, Dufresne, & Mestre, 1996),

In addition, to investigate the characteristics of thematic inference, the present study compared thematic inference with four other types of inference (i.e., goal, action, emotion, and

[2] The present study did not adopt a recognition task because while other types of inferred information (e.g., *The secretary married the accountant.*) can be falsely recognized in the task, thematic inference statements (e.g., *It was too late for Phil to prevent something that he had noticed.*) can hardly be recognized in the recognition task in the Implicit condition.

state), which are frequently generated in narrative passages. Although various classifications of inference have been proposed, as shown in section 2.2.2, inferences about a character's goals and actions were categorized as bridging inferences that construct relations between sentences in a text and therefore necessary for constructing coherence. On the other hand, inferences about emotion and state are classified as elaborative inferences not necessarily required for text comprehension, but lead to a deeper understanding. Previous studies demonstrated that readers generate more bridging inferences than elaborative inferences (e.g., Muramoto, 2000), and elaborative inference generation is a distinguishing feature of good and poor readers (e.g., Yoshida, 2003). The present study compared thematic inference with other inference types to investigate its distinctive characteristics. As few studies have directly examined thematic inference generation in L2 reading, the results of the present study are somewhat exploratory. However, considering that some studies categorize thematic inferences as elaborative (e.g., Whitney et al., 1991; Yoshida, 2003), it follows that thematic inferences should be generated less than goal and action inferences. On the other hand, given that thematic inference generation is necessary for a globally coherent representation of texts (Graesser et al., 1994), thematic inference would generate similar ratings as goal and action inferences. The research questions are as follows:

RQ1-1: Do Japanese EFL learners understand implicit themes in narrative texts by generating thematic inference?

RQ1-2: Does the generation of thematic inference differ from other inference types?

3.1.2 Method

3.1.2.1 Participants

The participants were 30 Japanese undergraduate students (11 female and 19 male) from

the same university. They were all Japanese EFL learners with different majors (e.g., engineering, international studies, medical science, sociology etc.) and had intermediate and advanced levels in English proficiency, having studied English for more than six years. Data of two participants who did not complete the given tasks were excluded from the analysis. Thus, results are based on the analysis of 28 participants.

3.1.2.2 Materials

Experimental passages

The materials provided were eight narrative passages by Seifert et al. (1986). One of the passages is shown in Table 3.1, and the others were similar to this story. All passages are presented in Appendix 1. The number of words, sentences, and the readability of each passage are provided in Table 3.2. As part of the experimental method, some low frequency words were simplified or given glossaries in Japanese. Each passage had an original title describing the theme of the passage in the form of an adage (e.g., *Closing the barn door after the horse is gone*, means to act too late to prevent something from happening).

Table 3.1

Sample of Experimental Passage in Experiment 1

Phil was in love with his secretary and was well aware that she wanted to marry him. However, Phil was afraid of responsibility, so he kept dating others and delayed proposing to her. Finally, his secretary got tired, began dating, and fell in love with an accountant. When Phil found out, he went to her and proposed marriage, showing her the ring he had bought. But by that time, his secretary was already planning her honeymoon with the accountant. <u>It was too late for Phil to prevent something that he had noticed.</u>

Note. The underlined sentence was only presented in the explicit version.

Table 3.2

Outline of Experimental Passages in Experiment 1

	Explicit				Implicit			
Texts	Word	Sentence	FRE	FKGL	Word	Sentence	FRE	FKGL
Alice	84	7	66.7	6.9	70	6	66.8	6.8
Bill	98	7	71.7	6.7	83	6	72.5	6.5
Brown	85	7	81	4.9	73	6	79.7	5.1
Burt	97	8	68.9	6.6	83	7	68.4	6.6
Ernie	105	7	58.6	8.8	92	6	55.1	9.3
Joe	100	7	73	6.6	86	6	76.2	6.1
Karen	94	6	66.7	7.4	78	5	62.3	7.9
Phil	92	6	64.3	8	79	5	61.2	8.6
M	94.71	7.00	69.51	6.84	80.71	6.00	68.71	6.90

Note. FKR = Flesch Reading Ease; FKGL = Flesch-Kincaid Grade Level. Values provided by Microsoft Word 2013 readability measurement tools.

Each experimental passage has a story structure in which the main characters' positive or negative action results in a corresponding outcome. In the Phil text, for example, the main character delayed marrying his secretary (negative action), which caused her to lose patience. When Phil eventually proposed, the secretary was already planning her honeymoon with someone else (negative outcome).

For experimental purposes, each passage had two versions: (a) explicit version, where the last sentence of the passage described the thematic statement and (b) implicit version where the sentence was deleted. In the Phil story, for example, the sentence, "It was too late for Phil to prevent something that he had noticed," was included as a thematic statement in the explicit

version. The thematic statement was created based on Seifert's et al. (1986) original title of the passage. The original title and thematic statement of each passage is shown in Appendix 1.

Inference verification task

Fourteen target statements for the inference verification task were created for each experimental passage. The details are as follows: (a) a *theme* statement, which describes the overall theme of the passage; (b) four *explicit* statements, which describe literal information (e.g., goal, action, emotion, and state) explicitly mentioned in the passage; (c) four *inference* statements, which describe information (e.g., goal, action, emotion, and state) not explicitly mentioned, but can be inferred from the passage; (d) five *inappropriate* statements, which describe information not mentioned or suggested in the passage. Examples of target statements in inference verification task are shown in Table 3.3.

It should be noted that theme statements were explicitly stated in the explicit version of each experimental passage, but not presented in the implicit version. Therefore, if participants generated thematic inferences in implicit passages and then encoded the inferences as part of the text memory, thematic statements are likely to be judged "yes" and evaluated as highly appropriate. Based on Muramoto (2000), all of the statements were presented to participants in L1 (Japanese) to avoid the effects of participants' surface text memory about word forms and sentence structure on the verification task.

Table 3.3

Examples of Target Statements in Inference Verification Task in Experiment 1

Theme		Phil が気付いた時には手遅れだった。[*It was too late for Phil to prevent something that he had noticed.*]
Explicit	Goal	秘書は Phil と結婚したかった。[*The secretary wanted to marry Phil.*]
	Action	Phil は指輪を買った。[*Phil bought a ring.*]
	Emotion	Phil の秘書は疲れてしまった。[*Phil's secretary got tired.*]
	State	Phil と秘書は恋をしていた。[*Phil was in love with his secretary.*]
Inference	Goal	Phil は秘書をとりもどそうとした。[*Phil tried to get his secretary back.*]
	Action	秘書は会計士と結婚した。[*The secretary married the accountant.*]
	Emotion	Phil はくやしい気持ちだった。[*Phil felt disappointed.*]
	State	Phil は優柔不断であった。[*Phil was indecisive*.]
Inappropriate	Goal	Phil は会計士と仲良くなりたかった。[*Phil wanted to get along with the accountant.*]
	Action	会計士は買い物に出かけた。[*The accountant went shopping.*]
	Emotion	会計士はとても悲しかった。[*The accountant felt very sad.*]
	State	秘書はお金持ちだった。[*The secretary was very rich.*]
	Theme	Phil が行動を起こすのは早すぎた。[*It was too early for Phil to take action.*]

Note. Target statements were presented to the participants in L1 (Japanese).

Before the experimental study, a norming study was conducted to verify the validity of inference statements. The participants were 10 graduate and undergraduate students who majored in English education. None participated in the experimental study. Participants read eight experimental passages of the implicit versions and rated whether each inference statement

could be inferred from the passage on a 5-point scale (1 = *cannot be inferred*, 5 = *can be inferred*). Additionally, when they rated the statement 4 or 5, they were required to classify it according to five categories (i.e., goal, action, emotion, state, and theme). The final versions of target sentences were determined based on the norming study (see Appendix 2).

3.1.2.3 Procedure

Participants were tested individually in sessions that lasted about 40 minutes. Eight passages were counterbalanced across two conditions (e.g., explicit and implicit) using a Latin square to ensure each participant read four passages in the explicit version and the remaining four passages in the implicit version. The experimental passages were presented in a random order to each participant. SuperLab 4.5 software (Cedrus, U.S.) were installed on a computer, and participants read passages using the Response Pad RB-730 (Cedrus, U.S.). The experimental phase followed the instructional and practice phases. In the instruction phase, the procedure was explained in Japanese; and in the practice phase, participants read a sample passage to confirm the procedure.

Figure 3.1. indicates the procedure of the inference verification task. Before viewing each passage, the signal *Ready?* appeared at the center of the screen and participants pushed the "yes" key to begin reading. They read each passage sentence-by-sentence in a self-paced fashion and pressed the "yes" key to signal that they had understood each sentence. They could not look back at prior sentences. The participants were asked to read carefully each sentence in order to complete the post-reading verification task; however, they were not given specific reading goals for inference generation. When the last sentence of a text disappeared from the screen, the target sentences for inference verification task appeared on the screen following the presence of a warning signal "***" for 1,000 milliseconds (ms). Participants were required to decide as soon as possible if the verification statement was appropriate, and could not refer back to the passage.

Each verification statement was presented on the screen until participants responded with the "yes" or "no" keys. This was followed by a 5-point scale (1: *not appropriate*; 5: *appropriate*), and the participants responded with the appropriate numeric keys. Fourteen statements for each passage were presented in random order, and the same procedure was repeated using eight passages.

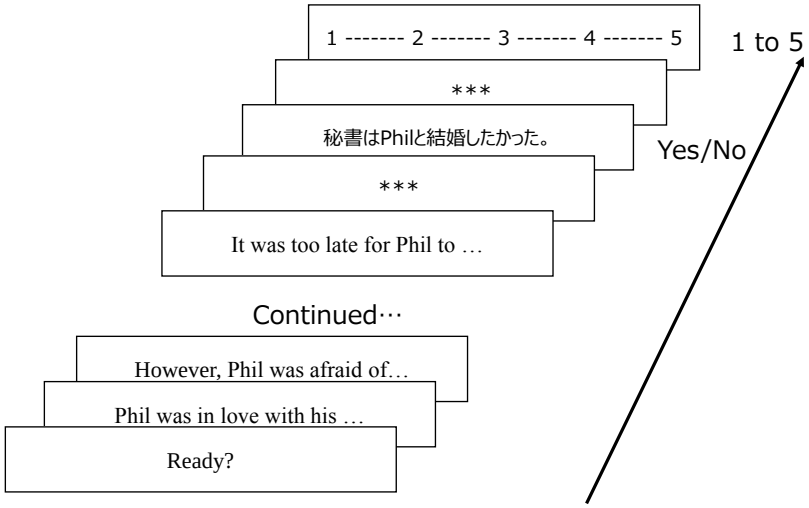

Figure 3.1. Procedure of the inference verification task in Experiment 1.

3.1.2.4 Scoring and Analysis

To examine whether participants generated the thematic inference in narrative passages, scoring comprised (a) the percentage of positive responses (i.e., "yes" responses) and then (b) where mean 5-point verification ratings were calculated and compared between two text conditions (i.e., explicit and implicit conditions). If participants generated thematic inference in implicit passages and then encoded the inferences as part of text memory, thematic statements were likely to be judged "yes" and evaluated as highly appropriate. As a result, the verification data of thematic statements should not differ between explicit and implicit conditions. In addition, to investigate the difference between thematic inference and the other four types of inference, scoring procedures (a) and (b) above for each inference statements (i.e., goal, action,

emotion, and state) in the implicit condition were calculated and compared.

3.1.3 Results

3.1.3.1 Thematic Statements in Explicit and Implicit Conditions

Table 3.4 shows the descriptive statistics of yes-response rates and 5-point scale ratings. Paired *t* tests were conducted on yes-response rate and 5-point scale respectively in order to determine whether there was significant differences between explicit and implicit conditions. The result showed that yes-response rates for thematic statements were significantly higher in Explicit than Implicit condition, $t(27) = 3.81$, $p = .001$, $d = .894$. Similarly, the 5-point scale rating data for thematic statements was significantly higher in Explicit condition than Implicit condition, $t(27) = 5.14$, $p < .001$, $d = 1.19$. These results indicated participants evaluated explicit themes as more valid than implicit themes, suggesting that understanding implicit themes by inference generation can be difficult for Japanese EFL learners.

Table 3.4

Yes-Response Rates and 5-Point Scale Ratings for Thematic Statements in Inference Verification Task in Experiment 1

	Yes-response rate			5-point scale		
Condition	*M*	95% CI	*SD*	*M*	95% CI	*SD*
Explicit	94.64	[90.77, 98.51]	10.45	4.59	[4.44, 4.74]	0.40
Implicit	82.14	[76.04, 88.24]	16.47	4.05	[3.87, 4.24]	0.49

3.1.3.2 Comparison Between Thematic Inference and Four Types of Inferences in Implicit Condition

Table 3.5 shows the descriptive statistics of yes-response rates and 5-point scale rating

for inference statements in the implicit condition. Separate one-way analysis of variances (ANOVAs) were conducted on yes-response rates and mean 5-point scale ratings respectively with Condition as a within-participant variable in order to determine whether there were significant differences among inference types. The results showed the main effects of Condition for the yes-response rates and the 5-point scale ratings (see Tables 3.6 and 3.7).

A post hoc comparison with Bonferroni correction (i.e., adjusted p value < .010) showed no significant differences in yes-response rates among the five types of inferences. On the other hand, the results of the 5-point scale ratings showed significantly higher ratings in goal and action inferences than emotion and state inferences (ps < .010). The rating of thematic inference was significantly higher than emotional inference, but similar to goal, action, and state inferences (see Figure 3.2).

Table 3.5

Yes-Response Rates and 5-Point Scale Ratings for Inference Statements in Implicit Condition in Experiment 1

	Yes-response rate			5-point scale		
	M	95% CI	*SD*	*M*	95% CI	*SD*
Theme	82.14	[76.04, 88.24]	16.47	4.05	[3.87, 4.24]	0.49
Goal	90.18	[84.36, 96.00]	15.72	4.38	[4.16, 4.59]	0.58
Action	84.82	[77.99, 91.65]	18.43	4.32	[4.12, 4.52]	0.54
Emotion	74.11	[66.35, 81.87]	20.95	3.62	[3.44, 3.80]	0.49
State	83.04	[76.34, 89.73]	18.07	3.88	[3.69, 4.06]	0.51

Table 3.6

Summary Table for One-Way ANOVA of the Effects of Inference Types on Yes-Response Rates in Experiment 1

Source	*SS*	*df*	*MS*	*F*	*p*	η_p^2
Inference Type	3767.86	4	941.96	3.64	.008	.119
Error (Inference Type)	27982.14	108	259.09			
Total	31750.00	112				

Table 3.7

Summary Table for One-Way ANOVA of the Effects of Inference Types on 5-Point Ratings in Experiment 1

Source	*SS*	*df*	*MS*	*F*	*p*	η_p^2
Inference Type	11.15	4	2.79	11.91	<. 001	.306
Error (Inference Type)	25.28	108	0.23			
Total	36.43	112				

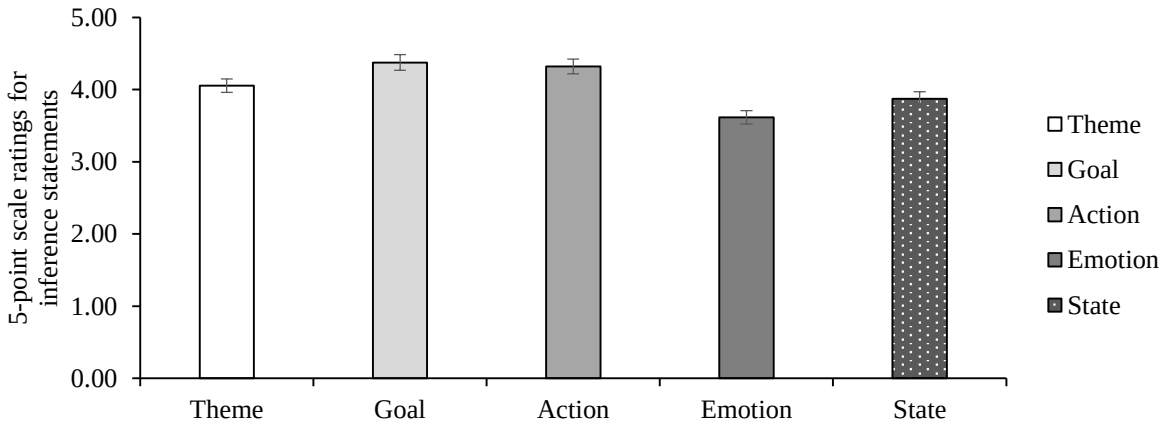

Figure 3.2. Five-point scale ratings for inference statements in Experiment 1.

3.1.4 Discussion

Do Japanese EFL learners understand implicit themes in narrative texts by generating thematic inference? (RQ1-1)

The results of thematic statements showed a significant difference between explicit and implicit conditions, which indicated that readers judged explicit themes more plausible than implicit themes. This result suggests that Japanese EFL learners are likely to have difficulty understanding implicit themes in narrative passages by inference generation.

As shown in section 2.2.2, although few researchers have directly investigated thematic inference in reading, some L1 reading theories, such as the minimalist hypothesis (McKoon & Ratcliff, 1992) and the constructionist theory (Graesser et al., 1994) have accounted for the generation of thematic inferences. Graesser et al. claimed that readers attempt to construct coherent meaning representations at both local and global levels. They predicted that thematic inferences are important for establishing global coherence of mental representations of narrative passages. However, they also stated that readers do not attempt globally coherent representations of the text under the following conditions: (a) the textual features lack global coherence and a message; (b) the reader does not have prerequisite background knowledge of the text; and (c) the reader has specific goals that do not require the construction of a meaningful representation (e.g., proofreading the text for spelling errors). When one or more of these conditions were present, the reader settles for local coherence or stops trying to achieve any coherence at all.

The reason participants did not evaluate implicit themes as highly as explicit themes in the present study can be explained by employing Graesser's et al. conditions above, even though, at first, all conditions appear to be satisfied. Regarding (a), because the experimental passages used in this study were organized around specific themes, text features were not likely to lack global coherence. As for (b), all the experimental materials were narrative passages about

everyday experiences involving characters' actions, goals, events, and emotions. Although the present study did not directly measure readers' previous knowledge about the passage, they were highly likely to understand these passages by adapting everyday experience rather than specific kinds of background knowledge. Regarding (c), the present study did not require the participants to engage in a specific goal that would hinder construction of global coherence because they knew that they would perform the verification task after reading each passage.

However, as the participants were told to perform the verification as a post-reading task, they might focus on comprehending each proposition and sentence in the passage rather than construct global coherence. Therefore, they probably set lower standards of coherence to achieve their reading goal for performing the verification task instead of building a globally meaningful coherence model. Indeed, the yes-response rate and 5-point scale verification was high for explicit statements. The yes-response rate was M = 83.71%, SD = 10.67 for explicit condition, and M = 84.82%, SD = 10.95 for implicit condition; the 5-point scale: M = 4.33, SD = 0.36 for explicit condition, and M = 4.33, SD = 0.28 for implicit condition. On the other hand, the score was low for inappropriate statements. The yes-response rate was M = 4.91%, SD = 5.73 for explicit condition, and M = 6.47%, SD = 7.88 for implicit condition; and the verification ratings were M = 1.61, SD = 0.48 for explicit condition, and M = 1.61, SD = 0.42 for implicit conditions. These results suggest that most readers succeeded in comprehending the passages at literal and textbase levels, which supports the possibility that readers set lower standards for coherence (i.e., understanding explicit or local information) rather than building global coherence).

The results of the present study are consistent with L1 empirical studies (e.g., Kurtz & Schober, 2001; Seifert et al., 1986) in that readers did not generate thematic inferences when specific reading goals were absent. For example, Seifert et al. (1986) showed that thematic processing occurred only when specific instructions to think about the themes in stories were

given before reading passages. They concluded that readers activated and encoded thematic information during reading narrative passages, but such encoding process depends on readers' strategies and given tasks. Similarly, Kurtz and Schober stated that the overall theme is not generated automatically, but is a later act of interpretation.

Additionally, the present study presented experimental passages in a sentence by sentence manner. This form of presentation might also inhibit readers to construct globally coherent situation models. According to the minimalist hypothesis, only a few inferences (e.g., referential, causal, antecedent) are automatically activated during reading—those based on quickly acquiring easily available information, and inferences required for local coherence of the text being read. Given that thematic inference generation requires readers to integrate widely separated pieces of textual information into mental representations, sentence-by-sentence text presentations might make it difficult for readers to integrate current with prior information given earlier in the passage. As shown by Morishima (2013), due to the limited cognitive resources available for discourse processing, maintaining global coherence of situation models is difficult, especially for L2 readers. Therefore, when readers can read the entire text instead of sentence-by-sentence text presentation, it might reduce the cognitive demands of reading comprehension and facilitate construction of globally coherent situation models.

Does the generation of thematic inference differ from other inference types? (RQ1-2)

The comparison of five types of inference statements showed that goal and action inferences were generated more easily than emotion and state inferences. This result is consistent with well-established reading theories (e.g., Graesser et al., 1994; McKoon & Ratcliff, 1992; van Dijk & Kintch, 1983) and L1 and L2 empirical studies (e.g., Muramoto, 2000; Yoshida, 2003). As stated in 2.2.2, goal and action inferences can be categorized into bridging inferences required for the construction of local text coherence. On the other hand,

emotion and state inferences are categorized into elaborative inferences not essential for coherence building between each textual segments, but contribute to global coherence of situation models. Consistent with Muramoto (2000), the present study supports the notion that readers generate bridging inferences more than elaborative inferences because bridging inferences are essential for text comprehension while elaborative inferences are not necessary for comprehension, but embellish what the text explicitly states. For example, "the secretary married the accountant" (action inference statement) can be inferred to bridge the gap between explicitly mentioned statements, "the secretary fell in love with an accountant," and "the secretary was already planning her honeymoon with the accountant." On the other hand, "Phil felt disappointed" (emotional inference statement) can be inferred from "his secretary was already planning her honeymoon with the accountant," but its necessity seems lower than the aforementioned action inference because it does not play a role in connecting textual information.

As for thematic inference generation, the 5-point scale verification indicated that thematic inference was significantly higher than emotional inference, but not different from goal, action, and state inferences. The descriptive statistics in Table 3.5 indicate that thematic inference is between bridging and elaborative inferences, which seems consistent with what reading theories assume. On the other hand, thematic inferences were rated higher than emotional inferences, and highly as goal and action inferences, suggesting that thematic inferences are obligatory rather than optional for discourse comprehension.

3.1.5 Conclusion of Experiment 1

The purpose of Experiment 1 was to examine, (a) whether Japanese EFL learners understand implicit themes in narrative texts by generating thematic inference, and (b) whether the generation of thematic inferences differ from other types of inference.

First, the comparison of theme-explicit and theme-implicit passages showed that learners had difficulty understanding implicit themes of narrative passages by generating inference (RQ1-1). In the present study, learners were not given a specific reading goal for constructing global coherence of the text; therefore, they might execute bottom-up processes such as encoding each textual information to construct local coherence rather than paying attention to top-down processing, such as comprehending the writer's messages conveyed by the overall passage.

Second, the present study also demonstrated that thematic inference was generated more than emotional inference, while generation of thematic inference was similar to goal, action, and state inferences (RQ1-2). It was suggested that while thematic inference was a kind of elaborative inference not necessary for comprehension, but it embellishes what the text explicitly stated. Thematic inference was generated as frequently as bridging inferences essential for text comprehension. On the other hand, regarding the findings for RQ1-1, Japanese EFL learners had difficulty generating thematic inferences in spite of its necessity in text comprehension. Therefore, further experiments should examine the effects of task instructions facilitating readers' global processing of thematic inference generation.

Although the present study showed that there was significant differences in the yes-response rate and the 5-point scale verification between explicit themes and implicit themes, it should be noted that the verification ratings for thematic inference statements for implicit and explicit conditions were relatively high (see Table 3.4). These results can be attributed to the fact that the inference verification task required participants to only judge presented statements instead of answering self-constructed themes. Consequently, the yes-response rate and the 5-point verification values were relatively high, even in the implicit conditions. Therefore, the next study adopts a thematic inference task in which readers answer an appropriate theme conveyed by the overall passage.

3.2 Experiment 2: Effects of Task Instructions on Thematic Inference Generation and Text Comprehension

3.2.1 Purpose and Research Questions

Experiment 2 was conducted in order to investigate the following two issues: (a) whether task instructions facilitate thematic inference generation among Japanese EFL learners and (b) whether task-induced strategic processing affects text comprehension.

The results of Experiment 1 suggested that Japanese EFL learners had difficulty understanding implicit themes in narrative passages by inference generation. Experiment 1 did not require readers to engage in a specific reading strategy; therefore, their reading goal might be to construct local coherence of the mental representation rather than building global coherence. Therefore, for learners to construct general themes in texts, some kinds of educational interventions are needed.

One possible intervention is giving task instructions to learners. As reviewed in 2.3, given a specific goal by task instructions, readers process a text from a particular viewpoint, and as a result, the importance of the information in the text changes (e.g., Bråten & Strømsø, 2009; Kaakinen & Hyönä, 2005; Kaakinen, Hyönä, & Keenan, 2002; Linderholm & van den Broek, 2002; Magliano et al., 1999; van den Broek et al., 2001). Considering the assumptions about theme comprehension and reading goals reviewed in Chapter 2, task instructions for thematic inference generation would induce learners to focus on important elements in the text (i.e., central action- and outcome-related information), which would facilitate building globally coherent representations of texts.

On the other hand, as reviewed in section 2.3.2, the effects of task instructions on text comprehension seem to be more complex and unstable for L2 and EFL reading than for L1 reading (e.g., Horiba, 2000, 2013; Yoshida, 2012). Therefore, one possibility here is that it will not facilitate text comprehension when too many resources are needed to achieve the given goal

(i.e., thematic inference generation). Another possibility is that the effects of task instructions interact with readers' L2 reading proficiency. Specifically, while high proficiency readers can control their reading processes more flexibly than low proficiency readers did; consequently, only high proficiency readers benefit from task instructions.

Experiment 2 examined the effects of task instructions on (a) strategic processing of thematic inference generation and (b) text comprehension by manipulating task instructions before reading passages and measuring participants' L2 reading proficiency. The research questions are as follows:

RQ2-1: Do Japanese EFL learners strategically generate thematic inferences when instructed to think about the overall message conveyed by the writer?

RQ2-2: Do task instructions aimed at thematic inferences affect Japanese EFL learners' text comprehension?

One methodological problem in Experiment 1 was that sentence-by-sentence text presentation might have prevented learners from constructing the global coherence of the text. To address this issue, Experiment 2 presents each experimental passage on one page at once so that participants can read passages with more natural reading settings. Moreover, Experiment 1 adopted an inference verification task—in which the participants answered whether and how presented statements were appropriate for the experimental passages—and did not measure the content of thematic inference generated by learners. Therefore, Experiment 2 adopted a thematic inference task (e.g., Kurtz & Schober, 2001), which directly measured whether and how learners constructed the theme of the passages. The details of this task are explained in the next section.

3.2.2 Method

3.2.2.1 Participants

A total of 64 undergraduate students (33 females, 31 males) participated in the present study. They were all Japanese EFL learners with majors in engineering, psychology, education, and literature. They have studied English for more than six years. The data of six participants who did not complete the given tasks were excluded from the analysis. Thus, the following analyses are based on the results of 58 participants.

3.2.2.2 Materials

L2 reading proficiency test

To assess the participants' English reading ability, a reading proficiency test was prepared. A total of five passages with 24 items were prepared. Counting the participants' proficiency level, the test included the pre-first (k = 5), second (k = 15), and pre-second (k = 4) grades in STEP test (Society for Testing English Proficiency, 1997, 2009). The items were all multiple choice questions, and the lengths of the passages were from 272 to 414 words.

Experimental passages

As experimental passages, four experimental passages (Phil, Burt, Ernie, Karen) were selected from the passages with implicit versions used in Experiment 1, with the support of seven graduate and undergraduate students majoring in English education. It was confirmed that these four passages implied specific moral points and they were similar in readability and topic familiarity. In contrast to Experiment 1, Experiment 2 presented these experimental passages on one page so that the participants could look at each passage at once on the paper.

Thematic inference task

As mentioned earlier, Experiment 2 adopted a thematic inference task in order to directly examine what kind of thematic inference was generated by Japanese EFL learners. In this task, the participants were asked to think about and write down the theme of each text in one sentence in their L1, Japanese.

3.2.2.3 Procedure

The participants were tested in a group setting (6 to 12 participants per group). The whole experimental section took about 65 minutes. First, the participants completed the reading proficiency test in 25 minutes. Then they were randomly assigned to one of the two reading conditions: reading for theme comprehension (the Task condition) or reading for text comprehension (the Control condition).

In the Task condition, the task instructions for theme comprehension on the basis of Seifert et al' s (1986) procedure was given to the participants before reading the texts. The participants were told that each text includes a different narrative theme that represents the writer's message or moral point, and that they were to think about the theme of the story as they read it. In addition, the researcher presented a well-known story "The Tortoise and the Hare" as an example (see section 1.1) to introduce the definition of theme. On the other hand, in the Control condition, participants were told to understand the passages carefully in order to comprehend texts; they were not told that each text includes an implicit theme.

Although the pre-reading instructions were different according to the reading condition, the participants in both conditions completed the following same procedure: (a) reading passages, (b) a written recall task, and (c) a thematic inference task. First, they read a passage silently for 90 seconds, and then completed an immediate written recall task in four and a half minutes. In the written recall task, the participants in both conditions were told to write down

all of what they remembered. They completed the recall task in Japanese, their native language, instead of English; because L2 or EFL readers may not be able to express the ideas that they actually comprehend in the target language due to constraints on their L2 writing skills (Lee, 1986). This procedure was repeated for each of the four passages and the order of presenting passages was counterbalanced.

After reading all texts and completing the recall tasks, participants completed a thematic inference task in four minutes. In this task, the participants were asked to write down the theme of each text in one sentence in Japanese using the first lines of each text as a cue. The definition and examples of theme was given to the participants before they started the thematic inference task.[3]

3.2.2.4 Scoring and Analysis

Written recall task

The recall production rate was used to indicate comprehension of explicit textual information. The experimental passages were divided into a set of idea units (IUs) on the basis of Ikeno's (1996) criteria. The standard of this division is as follows: (a) each idea unit consisted of a single clause (main or subordinate, including adverbial and relative clauses); (b) each infinitive, gerundive, and participle construction, nominalized verb phrase, and heavy adjunct (not complement) was also identified as separate idea units; and (c) argument and prediction conjuncts and disjuncts, such as train and/or bus, were separated into different idea units. This division was carried out by two raters including the researcher, and the agreement between them was 91.77%. The total number of IUs in each text was 19 to 20.

[3] This time limit of for each session was determined based on the pilot study with a small group of participants, and the participants in Experiment 2 also reported that they had finished reading each passage in the time limit.

Based on the idea units above, 30% of recall data were randomly selected and scored by two raters separately. In scoring the recall protocols, one point was given when an IU in the passages was correctly included in the recall protocols. The agreement between the two raters was 91.86%. Disagreements were resolved through discussion, and the remaining data were scored by the researcher alone. After scoring, the data were calculated in percentage form because the total number of IUs in each text was different.

In addition, in order to examine whether the learners can select information of central action and outcome which is necessary for thematic inference, IUs relating to the protagonist's main action and the outcomes were selected from four texts. The judgment was conducted by five raters including the researcher. Table 3.8 shows the action and outcome-related information of the four texts. On the basis of this judgment, the recall rates for each IU were compared between Control and Task conditions

Table 3.8

Central Action- and Outcome-Related IUs in Four Texts

Text	Central action		Outcome	
Burt	IU10	he started	IU16	Burt had qualified for
	IU11	reading about electronics	IU17	and found a better-paying job
	IU12	and decided	IU19	He also had more time
	IU13	to take courses by mail.	IU20	to enjoy himself
Ernie	IU10	He went to the shopping mall	IU18	Ernie was disappointed
	IU11	and looked for a dark blue security guard uniform,	IU19	that he had wasted money on uniforms
	IU12	and finally bought several		
Karen	IU9	the coach would warn the players	IU18	as he puffed heavily
	IU10	that they should avoid	IU19	on his long cigarette
	IU14	and especially smoking		
Phil	IU5	However, Phil was afraid of responsibility	IU19	his secretary was already planning her honeymoon
	IU7	dating others	IU20	with the accountant
	IU8	and delayed		
	IU9	proposing to her.		

Thematic inference task

For the thematic inference task, readers' answers were evaluated on the basis of the narrative themes of original texts (i.e., the titles of each text; see Appendix 1) and the dictionary

definition of each theme. First, each answer was categorized into (a) *correct*, which correctly described the writer's message conveyed through the overall passage; and (b) *incorrect*, which was inconsistent with suggested theme. Then, to conduct error analysis, inappropriate answers were further classified according to abstractness: *narrow* describing only a part of the passage rather than the overall passage and *broad* describing too abstract themes in the passage. Based on the criteria, 30% of the data were randomly selected and scored by two raters, and the inter-rater reliability was r = .856.

Table 3.9 shows the examples of readers' answers for Phil's text. The original theme is *closing the barn door after the horse is gone*, which means trying to take action when it is too late to prevent something from occurring. The narrow theme showed that the learner constructed the theme based on the local parts of the text "he kept dating others" rather than the overall text. On the other hand, the broad theme indicated that the learner were likely to understand the topic of the text (e.g., timing) but could not identify the specific message. The other example answers for each experimental passage are shown in Appendix 3.

Table 3.9

Examples of Participants' Answers for Thematic Inference Task in Phil's Text in Experiment 2

Category	Examples
Correct	It was already too late when he had noticed.
Incorrect	It is bad to date others. (Narrow)
	Timing is important. (Broad)

Note. Original answers were written in Japanese and translated in English by the author.

3.2.3 Results

3.2.3.1 L2 Reading Proficiency Test

The reliability of the reading proficiency test was acceptable (Cronbach's α = .83); therefore, the participants were divided into two proficiency groups (Upper, Lower) on the basis of this test. The number of participants and the mean scores are shown in Table 3.10.

Table 3.10

Descriptive Statistics of L2 Reading Proficiency Test in Experiment 2

	Proficiency	*n*	*M*	95% CI	*SD*
Control	Upper	14	19.57	[18.47, 20.67]	2.10
	Lower	14	11.79	[10.38, 13.20]	2.69
Task	Upper	15	19.93	[19.26, 20.60]	1.33
	Lower	15	13.87	[12.27, 15.47]	3.16

Note. Maximum possible score is 24.

In order to confirm whether the proficiency level of the two reading conditions was homogeneous, a 2 (Proficiency: Upper, Lower) × 2 (Condition: Control, Task) ANOVA was conducted on proficiency test scores. The results indicated that the significant main effect of Proficiency, $F(1, 53) = 115.04$, $p < .001$, $\eta_p^2 = .685$, whereas the main effect of Condition, $F(1, 53) = 21.158$, $p = .064$, $\eta_p^2 = .063$, and the interaction between these two factors, $F(1, 53) = 1.78$, $p = .188$, $\eta_p^2 = .033$, were not statistically significant. Therefore, these results confirmed that the proficiency level was approximately equal between the two reading conditions.

Table 3.11

Summary Table for Two-Way ANOVA of the Effects of Condition and Proficiency on the Score of L2 Reading Proficiency Test in Experiment 2

Source	*SS*	*df*	*MS*	*F*	*p*	η_p^2
Condition	1.67	1	1.67	21.16	.064	.063
Proficiency	703.76	1	703.76	115.04	< .001	.685
Condition × Proficiency	0.00	1	0.00	1.78	.188	.033
Error	320.97	53	5.73			
Total	16752.00	60				

3.2.3.2 Thematic Inference Task

To investigate RQ2-1, the data of the thematic inference task were scored and analyzed statistically. First, 232 answers (4 texts × 58 participants) were categorized into correct and incorrect answers. In the Control condition, 21 (18.75%) out of 112 answers were regarded as correct answers while 91 answers (81.25%) were incorrect. In the Task condition, on the other hand, 38 (31.67%) out of 120 answers were correct, while 82 answers (68.33%) were incorrect. In order to examine the relationship between task instructions and theme comprehension, a chi-square test was conducted. The results showed that the number of correct answers was significantly larger in the Task condition as compared to Control condition, $\chi^2(1) = 5.10$, $p = .024$, $\varphi = .148$.

Furthermore, an error analysis was conducted on incorrect answers in order to further investigate the effect of task instructions. In the Control condition, 63 (69.23%) out of 91 answers were categorized as narrow, while 28 answers (30.77%) were considered broad. In the Task condition, on the other hand, 40 (48.78%) out of 82 answers were narrow while 42 (51.23%) were broad. The results of a chi-square test indicated the task instructions were

significantly related to the content of error themes, $\chi^2(1) = 7.49$, $p = .006$, $\varphi = .208$, suggesting that while the participants in the Control condition were likely to answer narrow themes, which describe only a part of the text rather than the overall text, the participants in the Task condition were likely to answer broader themes, which did not capture the appropriate message conveyed by the writer.

3.2.3.3 Overall Recall of Explicit Textual Information

The recall production was analyzed in order to investigate the effects of reading goals on comprehension of explicit textual information. Table 3.12 shows the results of the mean recall productions for the four texts.

Table 3.12

Descriptive Statistics for the Percentage of Recall Production With Arcsine Transformation in Experiment 2

	Control				Task			
	n	*M*	95% CI	*SD*	*n*	*M*	95% CI	*SD*
Upper	14	51.09	[46.80, 55.38]	8.19	15	55.53	[53.19, 57.86]	4.61
Lower	14	38.73	[32.19, 45.26]	12.48	15	46.54	[42.32, 50.76]	8.34

A 2 (Condition: Control, Task) × 2 (Proficiency: Upper, Lower) two-way ANOVA was conducted on the mean recall production (see Table 3.13). The results showed that main effects of Condition, $F(1, 54) = 7.04$, $p = .010$, $\eta_p^2 = .115$, and Proficiency, $F(1, 54) = 21.39$, $p < .001$, $\eta_p^2 = .284$, were statistically significant. However, the interaction between the two factors was not statistically significant, $F(1, 54) = 0.53$, $p = .469$, $\eta_p^2 = .010$. As Figure 3.3 illustrates, the

participants in the Task condition better recalled than those in the Control condition did, and the participants with upper proficiency recalled better than those with lower proficiency did.

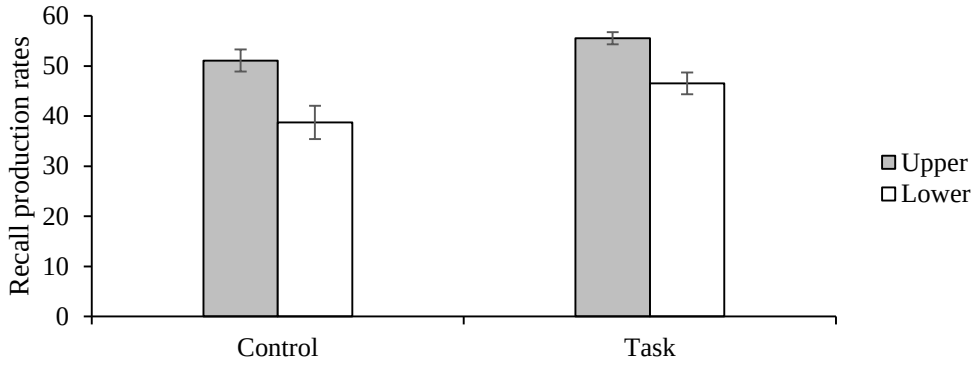

Figure 3.3. Percentage of recall production by reading condition in Experiment 2. Error bars represent standard errors.

Table 3.13

Summary Table for Two-Way ANOVA of the Effects of Reading Goal and Proficiency on the Recall Production Rates in Experiment 2

Source	*SS*	*df*	*MS*	*F*	*p*	η_p^2
Condition	543.54	1	543.54	7.04	.010	.115
Proficiency	1650.70	1	1650.70	21.39	.000	.284
Condition × Proficiency	41.07	1	41.07	0.53	.469	.010
Error	4166.77	54	77.16			

3.2.3.4 Recall Production Rates by Story Category

To investigate why the Task condition comprehended the text better than the Control condition, readers' recall protocols were further analyzed. Specifically, what information they

recalled better (and what information they did not) was examined. Recall production rates by story category (i.e. action, outcome, or others) are shown in Table 3.14.

Table 3.14

Recall Production Rates for Action, Outcome, and Other with Arcsine Transformation in Experiment 2

			Action		Outcome		Others	
		n	*M*	*SD*	*M*	*SD*	*M*	*SD*
Control	Upper	14	52.60	10.01	52.47	11.99	50.30	8.24
	Lower	14	40.84	10.30	34.68	17.15	39.44	12.18
	Total	28	46.72	11.63	43.58	17.11	44.87	11.60
Task	Upper	15	55.65	7.06	63.04	11.28	53.89	4.96
	Lower	15	46.20	12.72	56.04	8.97	45.02	9.01
	Total	30	50.92	11.19	59.54	10.63	49.46	8.45

A 3 (Category: Action, Outcome, Others) × 2 (Condition: Control, Task) × 2 (Proficiency: Upper, Lower) three-way ANOVA was conducted on the recall production rates (see Table 3.15 and Figure 3.4). The results showed significant main effects of Category, $F(2, 108) = 4.51$, $p = .013$, $\eta_p^2 = .077$. Condition, $F(1, 54) = 13.67$, $p = .001$, $\eta_p^2 = .202$. and Proficiency, $F(1, 54) = 24.08$, $p < .001$, $\eta_p^2 = .308$. Moreover, the interaction between Category and Condition was statistically significant, $F(2, 108) = 10.20$, $p < .001$, $\eta_p^2 = .150$. Follow-up tests with Bonferroni correction revealed that the simple main effect of Category was significant for the Task condition, indicating that Outcome information was recalled better than Action or Others (*p*s < .001). In addition, the simple main effect of Condition was also significant for Outcome information, indicating that the participants in the Task condition recalled Outcome

information better than those in the Control condition did ($p < .001$). These results suggest that when given a reading goal of theme comprehension, readers focus on outcome as well as action information, leading to the construction of more coherent mental representations.

Table 3.15

Summary Table for Three-Way ANOVA of the Effects of Category, Proficiency, and Condition on the Recall Production Rates in Experiment 2

Source	*SS*	*df*	*MS*	*F*	*p*	η_p^2
Within-participants						
Category	571.19	2	285.60	4.51	.013	.077
Category × Proficiency	49.10	2	24.55	0.39	.680	.007
Category × Condition	1293.65	2	646.83	10.20	.000	.159
Category × Proficiency × Condition	180.32	2	90.16	1.42	.246	.026
Error (Category)	6845.90	108	63.39			
Between-participants						
Proficiency	5211.99	1	5211.99	24.08	.000	.308
Condition	2959.78	1	2959.78	13.67	.001	.202
Proficiency × Condition	274.24	1	274.24	1.27	.265	.023
Error	11687.77	54	216.44			

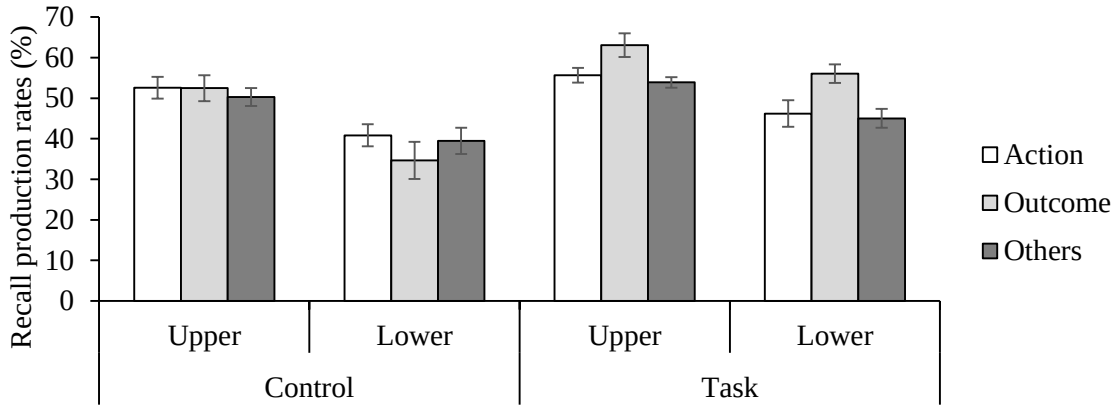

Figure 3.4. Recall production rates by story category in Experiment 2. Error bars represent standard errors.

3.2.4 Discussion

The effects of task instructions on thematic inference generation (RQ2-1)

The results of the thematic inference task indicated that the participants in the Task condition performed better in the thematic inference task than those in the Control condition, suggesting that Japanese EFL learners strategically generated thematic inferences when instructed to think about the overall message conveyed by the writer.

Some L1 previous studies also demonstrated that thematic inferences were not constructed without any specific task instructions (e.g., Kurtz & Schober, 2001; Seifert et al., 1986). In Experiment 1, the learners did not understand implicit themes better compared to when the narrative themes were explicitly mentioned in the passage. The possible reason that learners had difficulty generating thematic inference was that the learners might set their standards of coherence at textbase or local level when not given specific tasks. Experiment 2, which manipulated the participants' reading goal by giving task instructions, showed the positive effect of task instructions on thematic inference generation, which supported the

possibility suggested in Experiment 1. Therefore, it can be concluded that thematic inferences are strategic processing, which is sensitive to the reader's goals and comprehension strategies.

In addition, the error analysis of thematic inference task demonstrated that the content of the theme constructed by learners was related to the task instructions. Specifically, while the participants in the Control condition were likely to construct too narrow themes based on local and partial text information rather than the overall text, those in the Task condition were likely to construct too abstract themes. These results also supported the notion that thematic inference was task-induced strategic processing. Without specific instructions, learners paid too much attention to comprehending each proposition or constructing local coherence of the text; consequently, they constructed narrow themes that depended on explicitly stated information in the part of the text. In the Task condition, on the other hand, learners might pay less attention to processing each proposition in the text while they were likely to scan the text, resulting in the construction of too abstract themes. Thus, some learners did not take advantage of the given task. Similar findings were also found in Horiba (2013). In her study, the critique condition (i.e., reading a text by comparing the author's views with their own to evaluate them) was set in order to encourage higher level conceptual processing, which leads to the construction of stronger representations. Contrary to expectations, some readers were likely to set lower standards of coherence to achieve the reading goal and engaged in strategic processing at only a minimal level.

In sum, Experiment 2 confirmed that thematic inference was task-induced strategic processing rather than automatic processing with very little cognitive effort and processing resources. In addition, it was suggested that the task instructions adopted in Experiment 2 changed the allocation of cognitive resources during reading. Specifically, without any task instructions, learners tend to allocate much attention to lower-level linguistic processing to construct mental representation at the textbase level by engaging in a bottom-up manner. On

the other hand, task instructions facilitated top-down processing to extract the theme conveyed through the overall text. Magliano et al. (1999) demonstrated that when specific processing was facilitated by instructions, other types of processing were hindered in a trade-off manner. Therefore, Experiment 3 aimed to examine the effects of task instructions on processes during reading.

The effects of task instructions on text comprehension (RQ2-2)

The results of the recall task showed that the participants in the Task condition recalled better than those in the Control condition, which suggested that text comprehension was facilitated by task-induced strategic processing of thematic inference generation.

Some previous L2 studies showed that task instructions did not facilitate text comprehension (e.g., Horiba, 2013; Yoshida, 2012). The contradiction between the present study and these previous studies can be attributed to the general characteristics of task instructions. Yoshida (2012), for instance, gave three types of instructions: outlining, answering embedded questions, and reading only. In Horiba (2013), the participants were required to read for expression, for image, and for critique. In these studies, the readers were told to read a text to achieve their goals but they were not instructed to alter specific processing behaviors. On the other hand, some previous studies gave specific instructions, which facilitated text comprehension. Horiba (2000, Experiment 2) used the read-for-coherence instructions (i.e., pay attention to the relation between sentences), and Nahatame (2014) instructed learners to anticipate the outcome of the events described. The task instructions in Experiment 2 (i.e., to think about the theme of the story as they read it) were relatively specific, and the participants understood the goal of the task because they were given practice sessions. Therefore, the task instructions in the present study facilitated learners' text comprehension.

Additionally, the detailed analysis of recall production rates by information category showed that the readers in the Task condition recalled outcome-related information significantly better than those in the Control condition. It was demonstrated that the facilitative effect of task instructions were especially found in outcome-related information in the passages. Previous studies showed that without any specific reading goal, L1 and L2 readers tend to recall outcome statements better than action statements (e.g., Horiba et al., 1993). The Task condition in the present study also showed the presence of this pattern in learners' recall protocols. However, the most interesting finding of this study was that the recall rate for outcome information was significantly higher in the Task condition than in the Control condition. In the Task condition, readers needed to understand not only parts of the text but the text as a whole. Specifically, they needed to identify whether the outcome described in the text was consistent with the main character's actions (Dorfman & Brewer, 1994). Therefore, they focused on the key elements of action and outcome. In contrast, in the Control condition, the participants processed each type of information at the same level, leading to a lack of significant differences among story categories.

To sum up, task-induced strategic thematic inference facilitated Japanese EFL learners' text comprehension. It has been predicted that task instructions facilitating thematic inference might hinder text comprehension when too many resources are needed to achieve the given goal. However, the present study demonstrated that learners' text comprehension was not impaired by focusing attention on thematic inference generation. The specificity of task instructions given in the present study was relatively high, which allowed learners to alter strategic processing to achieve a given reading goal. Thematic inference is a kind of inference that contributes to global coherence of the text as the constructionist theory (Graesser et al., 1994) states. Therefore, strategic instructions for thematic inference generation helped learners to construct a globally meaningful representation.

Moreover, it had been assumed that the effects of task instructions interact with learners' L2 reading proficiency. Specifically, high proficiency readers can control their reading processes more flexibly than low proficiency readers; consequently, only high proficiency readers benefit from task instructions. However, the results showed only significant main effects of task instructions and proficiency, while interaction between these two factors was not found. It was suggested that task-induced thematic inference processing were appropriate educational intervention for both lower- and higher-proficiency learners. Experiment 3 further examines the effects of task instructions, focusing on the relationship between learners' proficiency and cognitive processes during reading.

3.2.5 Conclusion of Experiment 2

The purpose of Experiment 2 was to examine whether task instructions facilitate thematic inference generation and whether task-induced strategic processing affect text comprehension. To investigate these issues, Experiment 2 compared the Task condition (i.e., to read passages in order to understand the theme conveyed by the writer) and the Control condition (i.e., to read passages to comprehend).

First, the results of thematic inference task demonstrated that task instructions facilitated inference generation among Japanese EFL learners (RQ2-1). In addition, the error analysis showed that the learners were likely to construct too narrow themes without any specific instructions, while they were likely to construct abstract themes when given task instructions. Second, the results of the recall task showed that the participants in the Task condition recalled more text information, especially outcome-related information, than those in the Control condition. It was demonstrated that task-induced strategic processing facilitated text comprehension among Japanese EFL learners (RQ2-2).

Although the positive effects of task instructions were found in Experiment 2, the reason that such facilitation occurred was still unclear. The plausible interpretations are as follows. In the Control condition, learners tended to allocate too many cognitive resources to lower-level linguistic processing in order to construct mental representations at textbase and local level. As a result, they were likely to allocate fewer resources to higher-level conceptual processing for building a globally coherent representation of meaning, leading to poorer performance on thematic inference generation and text comprehension. As for the Task condition, on the other hand, learners not only engaged in understanding text information but also engaged in more conceptual processing such as connecting information and activating background knowledge in order to achieve their reading goals.

As Experiment 2 aimed to examine the effects of task instructions on thematic inference generation and text comprehension after reading passages (i.e., "products" of text comprehension), it did not directly focus on learners' allocation of cognitive resources during reading (i.e., "processes" of text comprehension). Therefore, the interpretations of the findings stated above need verification in the following experiment. To understand the effects of task instructions on text comprehension, further research is needed to investigate processes during reading as well as products of comprehension after reading (Horiba, 2013). Experiment 3 adopted a think-aloud method as a direct measure of learners' cognitive processes during reading comprehension.

3.3 Experiment 3: Effects of Task Instructions on Processes and Products of Narrative Comprehension

3.3.1 Purpose and Research Questions

Experiment 1 showed that Japanese EFL learners had difficulty understanding implicit themes through inference generation. On the other hand, Experiment 2 demonstrated that Japanese EFL learners strategically generated thematic inferences when instructed to think the overall theme conveyed by the writer. Moreover, such task instructions aimed at thematic inference facilitated learners' text comprehension. Thus, two experiments suggested that thematic inference was a kind of strategic processing facilitated by task-induced reading goals, and it contributed to building coherent and robust mental representations of texts.

While Experiment 2 demonstrated some important aspects of task instructions, it is necessary to examine the online processes that EFL readers engage in during reading in order to clarify the effects of task instructions. According to some researchers, distinguishing between the products and processes of reading comprehension is important because the mental representations constructed after reading texts (i.e., the products of reading comprehension) were constructed through moment-by-moment processes as the reader proceeds through the text (e.g., Kendeou, van den Broek, Helder, & Karlsson, 2014; Kintsch, 1998; McMaster, Espin, & van den Broek, 2014). Therefore, Experiment 3 gave the same task instructions to another group of Japanese EFL learners and measured their cognitive processes by adopting a think-aloud method.

A think-aloud method requires participants to verbalize whatever they are thinking while performing a task (e.g., Israel, 2015). This method has been widely used in L2/EFL reading research because reading is normally a silent, hidden process, and researchers cannot determine what is happening during reading by product-based assessment (e.g., Yoshida, 2008).

Regarding these advantage of the think-aloud method, the present study adopted this method in order to directly measure the cognitive processes that occurred during EFL reading.

In spite of the benefits, it should be noted that the think-aloud protocols cannot provide a full picture of processing (e.g., Yoshida, 2008). Even though the task instructions changed learners' standards of coherence or the goals for reading texts, they may not be reflected in the think-aloud protocols. Therefore, many researchers suggested that think-aloud data was interpreted by comparison and combination with other methodologies (e.g., Bowles, 2010; Magliano et al., 1999). The present study prepared a questionnaire in order to measure reading processes that learners tried to engage in during reading. The details are shown in the method section. A hypothesis (H) and two research questions of Experiment 3 are as follows:

H: The task instructions aimed at thematic inference generation facilitated text comprehension measured by the written recall task.

RQ3-1: Do task instructions aimed at thematic inference generation change EFL learners' reading goals measured by a questionnaire?

RQ3-2: Do task instructions aimed at thematic inference generation change EFL learners' cognitive processes measured by a think-aloud task?

Regarding the results of Experiment 2, it can be expected that EFL learners tend to allocate too many cognitive resources to lower-level linguistic processing (e.g., word and sentence analyses) and fewer resources to higher-level conceptual processing (e.g., inference generation, evaluation of the text content) when they were not give any specific task instructions. On the other hand, task instructions aimed at thematic inference generation can allow learners to engage in more conceptual processing in order to construct globally coherent mental representations of texts.

Although few studies examined the effects of task instructions aimed at thematic inference generation on reading processes among EFL learners, some previous L2 studies which adopted a think-aloud method provide meaningful insights into the present study. For example, Horiba (2000) demonstrated that while learners did not change their cognitive processes during reading according to the given task, the task instructions facilitated their recall production rates after they had finished reading. In contrast, Horiba (2013) showed that L2 learners' processes were partly changed by the task instructions, while the recall production rate did not differ between the Task conditions. Thus, the effects of task instructions on the processes and products of L2 reading comprehension are more complex and various than those of L1 reading comprehension. Therefore, according to the results of these studies, Japanese EFL learners' cognitive processes would change partly or would not change at all according to the task instructions.

Another possibility is that the task instructions cannot directly change the processes themselves, but learners' reading goal can be influenced by the task instructions. In this case, the effects of task instructions would not be found in the think-aloud data but in the reading strategy questionnaire.

3.3.2 Method

3.3.2.1 Participants

Participants were 30 Japanese undergraduate and graduate students (11 females and 19 males; range 18-24 years) majoring in a variety of fields, including social studies, engineering, biology, and medical science. All participants had studied English for more than six years as a part of their formal Japanese education, and their self-reported English proficiency levels ranged from intermediate to advanced. Participants were classified into two reading proficiency groups based on the L2 reading proficiency test.

3.3.2.2 Materials

Experimental passages

Four short narrative passages used in Experiment 2 were also used as the experimental passages in this study. In these passages, the themes were not explicitly stated.

Questionnaire

In order to test the effects of task instruction on the reading goals, participants were asked eight questions (shown in Table 3.16) that focused on the construction of coherent mental representations. These items were assumed to support the think-aloud data as they provided information related to whether or not learners tried to change their resource allocation during reading (even if differences between task instructions did not appear in the form of utterances during the think-aloud protocols).

The type of coherence that was assumed to be constructed during reading was addressed in questions one to five. These questions were based on Zwaan and colleagues' *event-indexing model*, which assumes that readers mentally represent five dimensions in narrative texts: time, space, causation, motivation, and protagonist (e.g., Zwaan et al., 1995; Zwaan & Radvansky, 1998). Questions six to eight assessed levels of coherence so that Q6 related to referential/anaphoric coherence, Q7 determined whether participants paid attention to the local coherence of the text, and Q8 focused on global coherence. All questions were presented in Japanese, and participants were asked to rate the degree to which they had paid attention to each of the eight perspectives during reading on a 5-point Likert scale (1: *not true* to 5: *true*).

Table 3.16

Items in Questionnaire About Reading Goals in Experiment 3

Q1	Character:	Who appeared in the story
Q2	Time:	The time course of the story
Q3	Space:	Where the events in the story took place
Q4	Cause:	Why the events in the story occurred
Q5	Goal:	What the characters in the story wanted to do
Q6	Anaphor:	What the pronouns (e.g., it, she, he, they) indicated
Q7	Local:	How information in the current sentence is related to prior or later sentences
Q8	Global:	How information in the current sentence is related to the overall text

Note. Questions were presented to participants in Japanese.

L2 reading proficiency test

In order to assess the participants' English reading ability, a 24-item reading proficiency test was prepared. The passages and questions were the same as in Experiment 2.

3.3.2.3 Procedure

The present study included the following four sections: (a) reading the four experimental passages with a think-aloud task, (b) a questionnaire relating to reading goals, (c) reading proficiency test (as an interference task), and (d) a written recall task. The entire test took approximately 90 minutes to complete.

Task instructions

In order to examine the effects of task instructions for thematic inference generation, the present study compared the condition with strategy instructions for thematic inference generation (the Task condition) with the condition without any specific instructions (the Control condition) as a within-participant factor. As in Experiment 2, in the Control condition, participants were only instructed to read the text for the purposes of completing the comprehension questions. In the Task condition, on the other hand, participants were asked to read the text in order to comprehend the author's message as conveyed through the overall text, and to report it aloud after completing the reading of each text (i.e., the thematic inference task). This instruction aimed to facilitate thematic inference generation during reading comprehension.

Although the main research interests of Experiment 3 relate to the think-alouds, and the written recall task, participants' answers to the thematic inference task indicated that the participants attempted to generate thematic inference in the Task condition according to the given task instructions. In the present study, to minimize task instruction effects other than those of thematic inference generation during reading, participants were not told that they would complete a written recall task.

Think-aloud task

A detailed explanation of the think-aloud task was first given in Japanese. The think-aloud task required participants to verbalize in their L1 whatever thoughts came to mind as they read each sentence. Whole passages were presented on a computer screen using SuperLab 4.5 (Cedrus, CA, USA). After reading the passage, participants were asked to press the enter key to move to the next passage. All verbal reports were recorded using an IC recorder. Before reading the experimental passages, participants read a practice passage that was similar in length to the experimental passages and which was provided without any specific task

instructions. After the practice session, participants completed the think-aloud task for two texts in the Control condition before doing the same for two texts in the Task condition. Assignment to the reading condition and the presentation order of the four texts was randomly counterbalanced.

Questionnaire

Participants were asked the eight questions shown in Table 3.16. As participants read the experimental passages, both in the Control and Task conditions, they answered the same questions twice. The specific procedure was as follows. First, the following instruction was given in Japanese: "Before you read the first two stories, you were asked to do so in order to answer post-reading comprehension questions. Under this condition, what did you pay attention to during the reading task? Please evaluate the following eight perspectives using a 5-point Likert scale." After judging their strategy use in the Control condition, participants were then given similar instructions: "Before you read the remaining two stories, you were asked to do so in order to consider the message conveyed by the author through the overall text. Under this condition, what did you pay attention to during the reading task?"

L2 Reading proficiency test

Before completing the written recall task, participants completed a reading proficiency test where they read five English passages and answered a total of 24 multiple-choice questions (approximate time = 25 min). As all participants in this study read two texts first in the Control condition and then two texts in the Task condition, this reading proficiency test was also intended to avoid any recency effect during the recall task.

Written recall task

After completing the reading proficiency test as an interference task, participants began a written recall task in which they wrote down, in Japanese, all of what they remembered. As they needed to recall four texts, the first sentence of each passage was presented as a recall cue. A time limit was not set in order to allow participants to recall as much of the texts as possible.

3.3.2.4 Scoring and Analysis

Think-aloud task

In order to score the think-aloud data, participants' verbal protocols were transcribed and then parsed into clauses. Horiba's (2013) framework was used for the categorization criteria, as this approach is intended for Japanese EFL learners and shares a similar theoretical interest to the present study (i.e., the effects of task instructions on learners' reading processes). Based on the purpose of the present study, some categories were combined and others were deleted. The author also included an additional category pertaining to thematic inference generation. Finally, each clause was categorized into one of the following 11 categories: (a) *word analysis*, (b) *sentence analysis*, (c) *paraphrase*, (d) *backward inference*, (e) *predictive inference*, (f) *thematic inference*, (g) *association*, (h) *evaluation*, (i) *reaction*, (j) *self-monitoring*, and (k) *other*. The definition of these categories are shown in Table 3.17, and the examples of protocols are shown in Appendix 4.

Table 3.17

Categories of Think-Aloud Protocols in Experiment 3

Process level	Category	Definition
Analysis	Word analysis Sentence analysis	The reader attempts to analyze the formal or semantic features of a word, phrase, and sentence, including L1 translation.
Paraphrase		The reader attempts to paraphrase the expression in the text to enhance his/her understanding.
Inference	Backward	The reader generates an inference that is intended to explain the contents of the current sentence by connecting it to prior text or on the basis of general knowledge.
	Predictive	The reader anticipates something about what will occur in the incoming text.
	Thematic	The reader states the main point or moral of the text.
Reader response	Association	The reader generates an inference that is brought to mind by the text that is not intended to enhance the understanding of the textual information.
	Evaluation	The reader makes a comment or states an opinion about the text that is evaluative.
	Reaction	The reader makes a comment to react, often emotionally, to the text.
Self-monitoring		The reader makes a comment about the degree of his/her own comprehension or use of a reading strategy.
Other		The reader comments on things that are not directly related to his/her comprehension of the text.

Both word and sentence analysis were subcategories of *analysis* where participants attempted to analyze the form or meaning of each piece of information at the surface memory level. In *paraphrase*, participants tried to construct their mental representations of the propositional textbase by paraphrasing the expression in the text in order to enhance his/her own understanding. Backward and predictive inferences were subcategories of *inference* where the reader engaged in relational and integrative processes based on the context of the passage or the reader's background knowledge, thus leading to the construction of a coherent representation of the text at the situational level. In *thematic inference*, participants tried to understand the author's messages or moral commentaries that were not explicitly stated in the text through the generation of inferences. On the other hand, *association*, *evaluation*, and *reaction* were considered as a *reader response*. Although a reader response itself is not intended to enhance the understanding of the text information, these categories can be regarded as strategies to actively understand an author's message through thinking about the relationships between the writer, the text, and the readers themselves. In regard to *self-monitoring*, participants commented on the degree of his/her comprehension or use of a reading strategy.

Two raters categorized one-fourth of participants' protocol data with an agreement rate of 85.15%. Disagreements were resolved through discussion, and the author scored any remaining data alone.

Written recall task

On the basis of the division of idea units conducted in Experiment 2, 30% of recall data were randomly selected and scored by two raters separately. In scoring the recall protocols, one point was given when an IU in the passages was correctly included in the recall protocols. The agreement between the two raters was 88.45%. Disagreements were resolved through

discussion, and the remaining data were scored by the researcher alone. The recall production rate was calculated and compared across reading conditions and L2 reading proficiency.

3.3.3 Results

3.3.3.1 L2 Reading Proficiency Test

The reliability of the reading proficiency test was acceptable (Cronbach's α = .88); therefore, the participants were divided into two proficiency groups (Upper, Lower) according to a median split of test scores. The number of participants and the mean scores are shown in Table 3.18. In order to confirm that there was a significant difference between two proficiency groups, a *t* test was conducted on the proficiency test scores. The result showed that there was a significant difference in test scores between these two groups $t(28) = 8.85$, $p < .001$, $d = 3.24$.

Table 3.18

Mean Scores on the L2 Reading Proficiency Test in Experiment 3

	M	95% CI	*SD*	*Max*	*Min*
Upper (*n* = 16)	20.88	[19.73, 22.02]	2.03	24	18
Lower (*n* = 14)	12.64	[10.81, 14.47]	3.03	16	6

3.3.3.2 Overall Recall Production Rates

To confirm that the task instructions aimed at thematic inference facilitated text comprehension, the recall production was analyzed again as in Experiment 2. Table 3.19 shows the results of the mean recall productions for the four texts. A 2 (Condition: Control, Task) × 2 (Proficiency: Upper, Lower) two-way ANOVA was conducted on the mean recall production (see Table 3.20). The main effect of Condition, $F(1, 28) = 6.27$, $p = .018$, $\eta_p^2 = .183$, was statistically significant. However, the main effect of Proficiency, $F(1, 28) = 1.94$, $p = .175$, η_p^2

= .065, and the interaction between the two factors, $F(1, 28) = 0.05$, $p = .823$, $\eta_p^2 = .002$, were not statistically significant. As Figure 3.5 illustrates, the participants in the Task condition better recalled than those in the Control condition did.

Table 3.19

Descriptive Statistics of the Percentage of Recall Production With Arcsine Transformation in Experiment 3

	Control			Task		
	M	95% CI	*SD*	*M*	95% CI	*SD*
Upper (*n* = 16)	43.41	[36.60, 50.22]	13.90	48.47	[45.04, 51.90]	7.00
Lower (*n* = 14)	37.16	[28.08, 46.23]	17.33	43.22	[37.28, 49.16]	11.34

Table 3.20

Summary Table for Two-Way ANOVA of the Effects of Reading Goal and Proficiency on the Recall Production Rates in Experiment 3

Source	*SS*	*df*	*MS*	*F*	*p*	η_p^2
Within-participants						
Condition	462.12	1	462.12	6.27	.018	.183
Condition × Proficiency	3.78	1	3.78	0.05	.823	.002
Error (Condition)	2062.98	28	73.68			
Between-participants						
Proficiency	493.98	1	493.98	1.94	.175	.065
Error	7144.05	28	255.15			

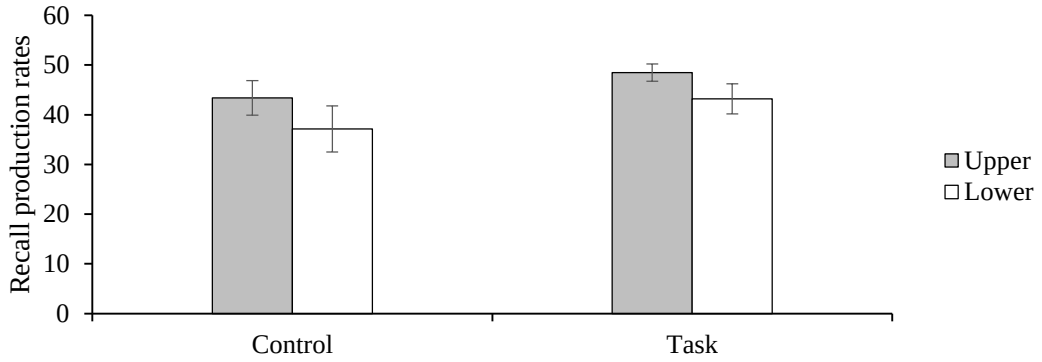

Figure 3.5. Percentage of recall production by reading condition in Experiment 3. Error bars represent standard errors.

3.3.3.3 Recall Production Rates by Story Category

To investigate the effects of task instructions on text comprehension in detail, learners' recall protocols were further analyzed as conducted in Experiment 2. Recall production rates by story category (i.e., action, outcome, and others) are shown in Table 3.21.

Table 3.21

Recall Production Rates for Each Story Category in Experiment 3

	Control			Task		
	Action	Outcome	Others	Action	Outcome	Others
Upper	49.82	54.50	39.90	53.66	64.38	43.38
(n = 16)	(19.68)	(26.51)	(14.98)	(12.78)	(16.93)	(11.26)
Lower	38.69	47.35	33.64	52.94	57.26	35.94
(n = 14)	(22.25)	(31.54)	(17.39)	(15.77)	(28.80)	(13.44)
Total	44.63	51.16	36.98	53.32	61.06	39.90
(N = 30)	(21.31)	(28.68)	(16.18)	(14.01)	(23.09)	(12.68)

Note. Standard deviations are in parentheses.

A 3 (Category: Action, Outcome, Others) × 2 (Condition: Control, Task) × 2 (Proficiency: Upper, Lower) three-way ANOVA was conducted on the recall production rates (see Table 3.22). The results showed significant main effects of Category, $F(2, 56) = 22.37$, $p < .001$, $\eta_p^2 = .444$, and Condition, $F(1, 28) = 6.51$, $p = .016$, $\eta_p^2 = .189$. On the other hand, the main effect of Proficiency, $F(1, 28) = 1.57$, $p = .221$, $\eta_p^2 = .053$, and the interaction between Category and Condition, which is statistically significant in Experiment 2, were not found, $F(2, 56) = 1.14$, $p = .327$, $\eta_p^2 = .039$. Follow-up tests with Bonferroni correction revealed that Action and Outcome were recalled better than Others (ps $< .001$) while there was no significant difference between Action and Outcome ($p = .069$).

These results suggest that task instructions facilitated recall production of each story category, regardless of L2 reading proficiency. Although the interaction between task instructions and information category was not found, the descriptive statistics showed that the effects of task instructions were appeared more in the recall production of Action and Outcome information compared to Others (see Figure 3.6). These results suggested that learners were likely to focus on important elements for thematic inference generation when given task instructions.

Table 3.22

Summary Table for Three-Way ANOVA of the Effects of Task Instructions, Information Type, and Proficiency on the Recall Production Rates in Experiment 3

Source	*SS*	*df*	*MS*	*F*	*p*	η_p^2
Within-participants						
Condition	2372.092	1	2372.092	6.51	.016	.189
Condition × Proficiency	106.623	1	106.623	0.29	.593	.010
Error (Condition)	10202.781	28	364.385			
Category	9433.990	2	4716.995	22.37	.000	.444
Category × Proficiency	11.954	2	5.977	0.03	.972	.001
Error (Category)	11807.741	56	210.853			
Condition × Category	436.444	2	218.222	1.14	.327	.039
Condition × Category × Proficiency	303.500	2	151.750	0.79	.458	.028
Error (Condition × Category)	10725.707	56	191.530			
Between-participants						
Proficiency	1973.553	1	1973.553	1.57	.221	.053
Error	35300.140	28	1260.719			

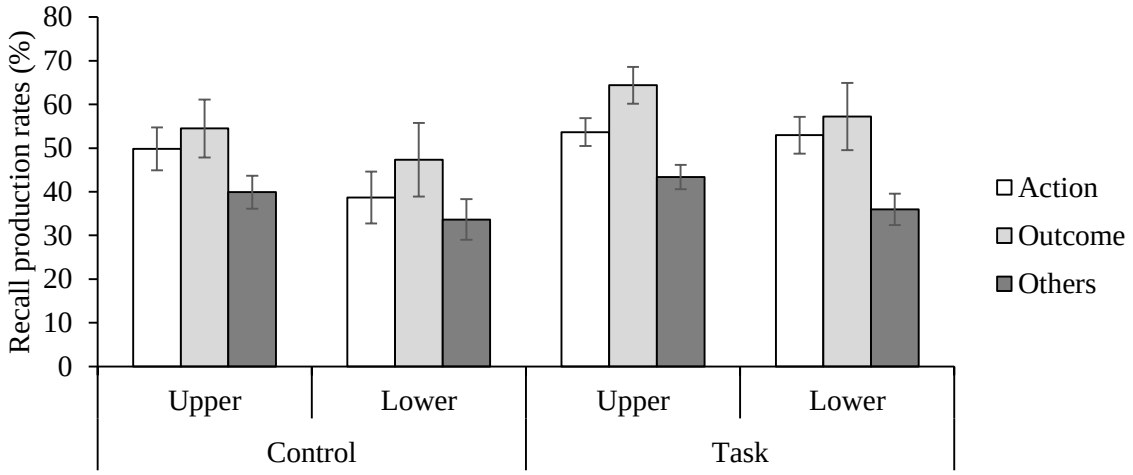

Figure 3.6. Recall production rates by story category in Experiment 3. Error bars represent standard errors.

3.3.3.4 Questionnaire

The questionnaire aimed to investigate the effects of task instructions on learners' reading goals to read the passage; therefore, they were asked to evaluate what they paid attention to during reading from the perspectives of construction of situation models of texts. The results of the questionnaire are presented in Table 3.23.

A 2 (Proficiency: Upper, Lower) × 2 (Condition: Control, Task) two-way MANOVA was conducted on 5-point scale judgment scores (see Table 3.24). The results showed that while the main effect of Proficiency and the interaction between two factors were not significant, the main effect of task instructions were significant, $F(8, 21) = 8.38$, $p < .001$, $\eta_p^2 = .761$.

Table 3.23

Reading Goals by Proficiency and Task Condition in Experiment 3

		Upper		Lower	
Item	Perspectives	Control	Task	Control	Task
Q1	Character	3.88 (1.09)	3.31 (1.35)	3.86 (1.29)	2.93 (1.38)
Q2	Time	3.31 (1.20)	3.13 (1.36)	3.86 (1.23)	3.50 (1.02)
Q3	Space	2.69 (1.49)	3.13 (1.50)	3.29 (1.33)	3.00 (1.30)
Q4	Cause	3.75 (1.00)	4.63 (0.62)	3.93 (1.21)	4.50 (1.09)
Q5	Goal	3.56 (1.31)	4.69 (0.48)	3.14 (1.46)	4.07 (1.27)
Q6	Anaphor	3.50 (1.59)	3.44 (1.41)	3.50 (1.29)	3.21 (1.48)
Q7	Local	3.56 (1.36)	3.63 (1.20)	3.71 (1.27)	3.79 (1.25)
Q8	Global	3.25 (1.44)	4.19 (1.11)	3.36 (1.22)	4.00 (1.18)

Note. Standard deviations are in parentheses.

Table 3.24

Summary Table for Two-Way MANOVA of the Effects of Condition and Proficiency on Reading Goal in Experiment 3

Source	Pillai's Trace	*F*	*df*	*Error df*	*p*	η_p^2
Between-participants						
Proficiency	0.16	0.51	8	21	.834	.163
Within-participants						
Condition	0.76	8.38	8	21	.000	.761
Condition × Proficiency	0.17	0.53	8	21	.820	.168

Follow-up ANOVAs with Bonferroni correction were conducted on rating scores for each question to examine whether participants' reading goals were changed according to the task instructions (see Table 3.25). The results indicated that the main effect of Condition was significant for Character ($p < .001$), Causal ($p < .001$), Intentional ($p = .001$), and Global ($p = .004$), while other perspectives were not different between the Control and Task conditions. As shown in Figure 3.7, the rating score on Character was higher in the Control condition than the Task condition, while the scores on Causal, Intentional, and Global were higher in the Task condition than the Control condition.

Table 3.25

Summary Table for Follow-Up One-Way ANOVAs of the Effects of Condition on Reading Goal in Experiment 3

Source		*SS*	*df*	*MS*	*F*	*p*	η_p^2
Condition	Character	8.30	1	8.30	17.30	.000	.382
	Time	1.12	1	1.12	1.36	.254	.046
	Space	0.09	1	0.09	0.11	.740	.004
	Causal	7.81	1	7.81	17.37	.000	.383
	Intentional	15.74	1	15.74	15.03	.001	.349
	Referential	0.45	1	0.45	0.98	.330	.034
	Local	0.07	1	0.07	0.09	.767	.003
	Global	9.32	1	9.32	10.01	.004	.263

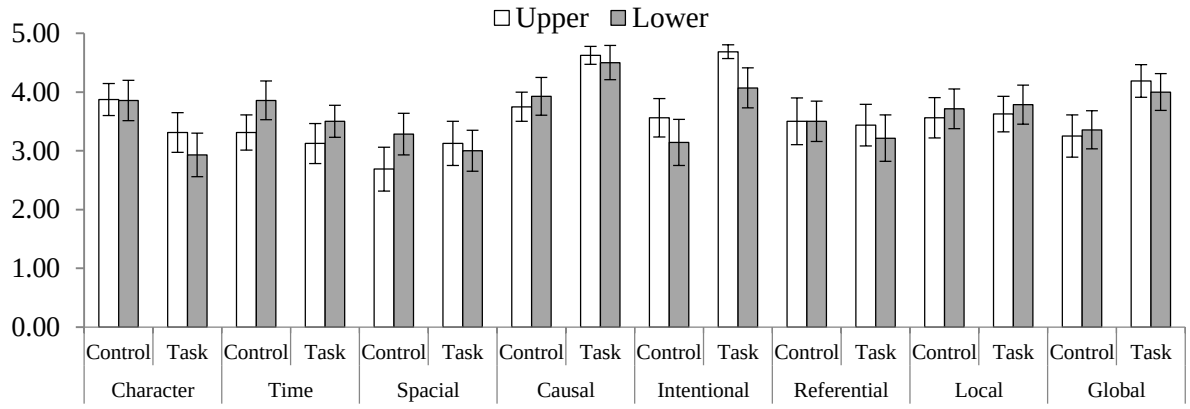

Figure 3.7. Rating scores for questionnaire in Experiment 3. Error bars represent standard errors.

These results suggested that the learners tried to pay much attention to causal relations, character's intention, and global coherence of the narrative texts; on the other hand, less attention as paid to who actually appeared in the stories. As the main effect of Proficiency and the interaction between two factors were not significant, the patterns of task effect were likely to be similar between proficient learners and less proficient learners.

3.3.3.5 Think-Aloud Protocols

Table 3.26 shows the number of think-aloud comments for each category by Task condition and L2 reading proficiency. As all the participants read two experimental passages for each Task condition, the values in this table indicate the average of two passages. The total number of comments were M = 36.72 in the Control condition and M = 37.13 in the Task condition. Although these values seem smaller compared to other previous studies adopting the think-aloud task (e.g., Horiba, 2000, 2013), it can be attributed to the length of experimental passages used in the present study.

Table 3.26

Number of Think-Aloud Comments for Process Level and Category by Task Condition and Proficiency in Experiment 3

		Control				Task			
		Upper		Lower		Upper		Lower	
Process	Category	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Analysis	Word	6.50	4.31	12.14	5.38	4.94	5.12	10.79	6.33
	Sentence	16.63	5.43	18.93	5.53	15.00	5.29	17.43	4.36
Paraphrase		6.06	4.09	4.00	1.71	6.88	3.22	7.29	3.71
Inference	Backward	1.25	1.84	0.50	0.76	2.19	1.42	0.57	0.85
	Predictive	0.13	0.34	0.00	0.00	0.38	0.62	0.00	0.00
	Thematic	0.00	0.00	0.00	0.00	1.31	2.27	0.07	0.27
Response	Association	0.44	0.89	0.07	0.27	0.31	0.48	0.21	0.58
	Evaluation	0.19	0.54	0.00	0.00	0.19	0.40	0.07	0.27
	Reaction	0.63	1.15	0.29	0.61	0.75	1.53	0.29	0.47
Monitoring		2.44	2.78	2.71	2.49	2.44	2.80	3.50	2.93
Other		0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.18
Total		34.25	11.76	38.71	8.62	34.38	12.45	40.29	9.86

To examine the cognitive resource allocation to each process level, the proportion of think-aloud comments were calculated per six categories (i.e., Thematic, Analysis, Paraphrase, Inference, Response, Monitoring) by Task condition and reading proficiency (see Table 3.27). The categorization was basically based on Horiba (2013), but the following points should be noted. First, although *thematic inference* was essentially a subcomponent of Inference, it was categorized as an independent category because thematic inference generation was the main

interest of the present study. Second, the present study identified distinguished Analysis (i.e., analyzing and translating words and sentences) and Paraphrase (i.e., paraphrasing the expression in the text to enhance their own understanding) in order to examine the effects of task instructions on cognitive resource allocation in detail by differentiating the levels of surface memory and propositional textbase.

Table 3.27

Proportion of Think-Aloud Comments for Process Level by Task Condition in Experiment 3

	Control				Task			
	Upper		Lower		Upper		Lower	
Process	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Thematic	0.00	0.00	0.00	0.00	4.60	7.61	0.26	0.95
Analysis	66.70	15.59	79.97	8.10	56.94	12.98	69.78	12.02
Paraphrase	17.30	10.53	10.60	4.94	19.88	7.40	18.23	9.70
Inference	4.25	6.38	1.28	2.03	7.16	3.53	1.40	2.52
Response	5.09	10.64	0.94	1.89	4.41	7.03	1.48	2.52
Monitoring	6.66	8.89	7.04	7.40	7.02	7.54	8.69	6.38

To examine the effects of task instructions and L2 reading proficiency on resources allocation during reading, a 2 (Condition: Control, Task) × 2 (Proficiency: Upper, Lower) MANOVA was conducted on the proportion of think-aloud comments, with six process-level categories as repetition variables (see Table 3.28). The results showed that the main effects of Condition, $F(6, 23) = 6.43$, $p < .001$, and Proficiency, $F(6, 23) = 3.49$, $p = .013$, and the interaction between two factors, $F(6, 23) = 2.62$, $p = .044$, were statistically significant respectively.

Table 3.28

Summary Table for Two-Way MANOVA of the Effects of Condition and Proficiency on Proportion of Think-Aloud Comments in Experiment 3

Source	Pillai's Trace	*F*	*df*	*Error df*	*p*	η_p^2
Between-participants						
Proficiency	.48	3.49	6	23	.013	.477
Within-Participants						
Condition	.63	6.43	6	23	.000	.626
Condition × Proficiency	.41	2.62	6	23	.044	.406

Table 3.29 shows the results of follow-up ANOVAs conducted on the proportion of think-aloud comments for each category. Follow-up tests with Bonferroni correction (i.e., adjusted *p* value < .008) indicated that the participants in the Control condition produced significantly more comments on Analysis than those in the Task condition, $F(1, 28) = 23.82$, $p < .001$, $\eta_p^2 = .460$. Moreover, while learners in Upper group produced more comments on Inference than those in Lower group, $F(1, 28) = 16.39$, $p < .001$, $\eta_p^2 = .369$, the comments on Analysis were produced more in the Lower group than the Upper group, $F(1, 28) = 9.96$, $p = .004$, $\eta_p^2 = .262$. As for other process levels, either significant main effects or the interactions between two factors were not found (*p*s > .008). These results are discussed in discussion section in detail.

Table 3.29

Summary Table for Follow-Up One-Way ANOVAs on the Proportion of Think-Aloud Comments in Experiment 3

Source		*SS*	*df*	*MS*	*F*	*p*	η_p^2
		Within-participants					
Condition	Thematic	87.90	1	87.90	5.58	.025	.166
	Analysis	1485.38	1	1485.38	23.82	.000	.460
	Paraphrase	388.81	1	388.81	5.55	.026	.165
	Inference	34.23	1	34.23	2.14	.154	.071
	Response	0.07	1	0.07	0.00	.946	.000
	Monitoring	15.09	1	15.09	0.58	.452	.020
Condition × Proficiency	Thematic	70.39	1	70.39	4.47	.044	.138
	Analysis	0.68	1	0.68	0.01	.917	.000
	Paraphrase	95.31	1	95.31	1.36	.253	.046
	Inference	29.20	1	29.20	1.83	.187	.061
	Response	5.56	1	5.57	0.35	.560	.012
	Monitoring	6.34	1	6.35	0.24	.625	.009
Error (Condition)	Thematic	440.80	28	15.74			
	Analysis	1746.04	28	62.36			
	Paraphrase	1961.75	28	70.06			
	Inference	447.32	28	15.98			
	Response	447.37	28	15.98			
	Monitoring	725.90	28	25.93			

		Between-participants					
Proficiency	Thematic	70.39	1	70.39	4.47	.044	.138
	Analysis	2546.12	1	2546.14	9.96	.004	.262
	Paraphrase	260.77	1	260.77	3.54	.070	.112
	Inference	284.26	1	284.26	16.39	.000	.369
	Response	187.62	1	187.62	2.48	.127	.081
	Monitoring	15.70	1	15.70	0.17	.681	.006
Error	Thematic	440.80	28	15.74			
	Analysis	7158.49	28	255.66			
	Paraphrase	2062.37	28	73.66			
	Inference	485.54	28	17.34			
	Response	2120.87	28	75.75			
	Monitoring	2554.86	28	91.25			

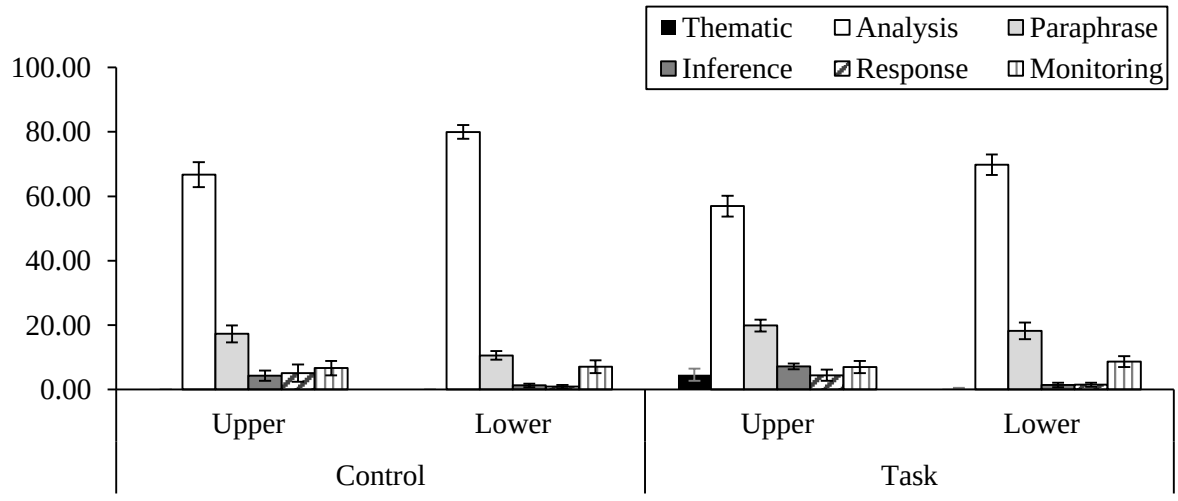

Figure 3.8. The proportion of think-aloud comments by task condition and L2 reading proficiency in Experiment 3. Error bars represent standard errors.

3.3.3.6 Relationship Between Learners' L2 Reading Proficiency and Cognitive Processes During Reading

To further examine the relationships between L2 reading proficiency and the allocation of cognitive resources, correlation analysis was conducted on the scores of reading proficiency test and the proportion of each think-aloud comments by Task conditions. The overview of results is shown in Table 3.30.

Table 3.30

Correlations Between L2 Reading Proficiency and Cognitive Resource Allocation During Reading by Task Conditions in Experiment 3

	L2 Reading Proficiency			
	Control		Task	
Process level	*r*	*p*	*r*	*p*
Thematic	N.A	N.A	.340	.066
Analysis	–.370*	.044	–.337	.069
Paraphrase	.295	.113	.083	.661
Inference	.217	.250	.611**	.000
Response	.315	.089	.170	.370
Monitoring	–.102	.593	–.205	.277

Note. NA = not applicable. Comments on Thematic in the Control condition were not found.

In the Control condition, the results revealed that L2 reading proficiency had medium negative correlation with the proportion of comments on Analysis (r = –.370, p = .044), suggesting that the lower the learners' proficiency was, the more they analyzed each word and sentence in the passages. Also in the Task condition, a similar pattern was found (r = –.337, p

= .069). More importantly, while L2 reading proficiency had positive correlation with the proportion of comments on Inference in the Task condition (r = .611, p < .001), such a pattern was not found in the Control condition (r = .217, p = .250). These results suggested that when proficient learners were given the task instructions, they tended to generate more backward and predictive inferences. In sum, the lower proficient learners were likely to allocate more cognitive resources to lower-level linguistic processing (i.e., analyzing words and sentences) regardless of the task instructions. On the other hand, the task instructions allowed proficient learners to allocate more cognitive resources to higher-level conceptual processing (i.e., backward and predictive inferences). These data were further discussed in section 3.3.4.

3.3.4 Discussion

Effects of task instructions aimed at thematic inference generation on text comprehension (Hypothesis 1)

The results of the written recall task demonstrated that the learners recalled more textual information in the Task condition than in the Control condition. The result was consistent with the recall production rate in Experiment 2, which confirmed that the task instructions aimed at thematic inference facilitated Japanese EFL learners' text comprehension.

On the other hand, different trends were found in the Task condition and the Control condition as follows. While significant interaction between Condition and Category was found in Experiment 2 (see Figure 3.4), such interaction was not found in Experiment 3 (Figure 3.6). Specifically, in Experiment 2, while the learners in the Control condition recalled action, outcome, and others at the same level, those in the Task condition recalled outcome-related information more than action and others. In Experiment 3, on the other hand, outcome- and action-related information were recalled better than others regardless of the task instructions. Such a difference in recall production patterns between Experiments 2 and 3 might be attributed

to the time limit for reading texts. In Experiment 3, the participants read experimental passages at their own pace with think-aloud tasks, whereas the participants in Experiment 2 read each passage for around 90 seconds. This time limit of Experiment 2 was determined based on a pilot study conducted with a small group of participants. The participants in Experiment 2 reported that they finished reading each passage in the time limit. However, when the participants were required to read the passage with time limit, they needed to allocate their attention more selectively than when they had no time limit for reading passages. As a result, the interaction between the task instructions and what they recalled could be more clearly found in Experiment 2 than Experiment 3.

Although such different trends were found in two experiments, the descriptive statistics in Experiment 3 showed that the effects of task instructions seemed to appear more clearly in Action and Outcome information, which is especially important in theme comprehension (e.g., Dorfman & Brewer, 1994; Zhang & Hoosain, 2001, 2005) than in Others. Therefore, it can be concluded that the task instructions aimed at thematic inference generation helped Japanese EFL learners pay attention to important elements, which contributed to construct a globally coherent and robust representations of narrative texts.

According to the notion of *standards of coherence* (van den Broek et al., 2001; van den Broek et al., 1995), the reason that the task instructions affected the products of text comprehension (i.e., learners' performance in the recall task) was that readers set their reading goals and criteria for comprehension based on the reason or purpose for reading by given task instructions. They set stronger or weaker criteria for how well they must comprehend a text, and this consequently changed the kinds of cognitive processes they actually engaged in. Therefore, the following discussions focused on why such facilitative effects were caused by the task instructions, from the perspective of reading goals and cognitive resource allocation during reading.

Effects of task instructions aimed at thematic inference generation on Japanese EFL learners' reading goal (RQ3-1)

The results of the questionnaire indicated that the participants paid more attention to information about characters in the Control condition than in the Task condition. Although character information is essential to comprehend narrative texts, it is less important for thematic inference generation because implicit themes can be constructed without specific and concrete information about characters. In many cases, for example, what the name of the character is, how many characters are in the story, and the appearance of the characters do not matter in making thematic inference. As a result, when learners were required to generate thematic inferences by task instructions, they paid less attention to such character information.

In contrast, they paid much attention to global coherence of texts when they were given the task instructions aimed at thematic inference generation. As thematic inference is a kind of inferences that contribute to building global coherence of narratives (e.g., Graesser et al., 1994), the learners' attention was strategically shifted to building global coherence rather than local coherence. Moreover, the task instructions allowed learners to pay attention to causal relationships and characters' intention. Among the five situational dimensions assumed by an event-indexing model (e.g., Zwaan & Radvansky, 1998), causal and intentional dimensions were related to connecting several events rather than single event. Therefore, they have been regarded as important dimensions in order to organize overall and global text representations. Regarding this assumption, it is reasonable that the learners paid more attention to these elements in the Task condition than in the Control condition.

Moreover, given that recall production rate in the Task condition was significantly higher than that in the Control condition, it is plausible that they paid more attention to characters' intention. According to the notion of a *goal–attempt–outcome episode* introduced by Trabasso and Wiley (2005), goal information (i.e., characters' intentions) plays a central role in narrative

stories because character's action is motivated by the goal, and the story outcomes depend on whether the goal is achieved or not. Therefore, the facilitative effects of task instructions on recall production might be attributed to the fact that the learners allocated much attention to characters' intention, which leads to building coherent and well-organized representations of narrative texts. In sum, the task instructions aimed at thematic inference generation affected Japanese EFL learners' reading goal. When the learners were required to construct implicit themes of texts, their attention was directed toward causal relationships, character's intentions, and global coherence of texts. The results of questionnaire showed no significant main effect of L2 reading proficiency and the interaction between proficiency and task instructions, suggesting that the effects of task instructions on learners' reading goals did not differ between two proficiency groups.

Effects of task instructions aimed at thematic inference generation on Japanese EFL learners' cognitive processes during reading (RQ3-2)

As a whole, the results of think-aloud data revealed that the effects of task instructions appeared only in a part of reading processes. Similar findings were observed in previous studies with L2/EFL learners (e.g., Horiba, 1996, 2000). Japanese EFL learners had difficulty controlling their processes flexibly because lower-level linguistic processing was less automatized.

Nevertheless, the task instructions influenced Japanese EFL learners' allocation of cognitive resources during reading. The following discussion especially focused on the following cognitive processes: (a) thematic inference, (b) word and sentence analysis, and (c) backward and predictive inferences.

First, as for thematic inference, which the task instructions in the present study aimed to facilitate, the descriptive statistics in Table 3.26 and Table 3.27 showed that only a small

number of comments were produced even in the Task condition. This small number of comments on thematic inference might be attributed to the characteristics of thematic inference itself. Thematic inference is triggered by lengthy stretches of text, rather than single words, clauses, or sentences; therefore, readers conclude the overall theme only at the story's end in many cases (Kurtz & Schober, 2001). Therefore, learners might have avoided producing comments until the post-reading thematic inference tasks given after reading passages.

Indeed, the number of participants who produced comments on thematic inference was only 8 out of 30 (27%) in the Task condition. Furthermore, only two participants produced comments on themes in the middle of the experimental passages. These two learners used strategies where they surmised a possible theme at the middle of the passage and revised it into a more appropriate theme according to the textual information given later. The following think-aloud comments are examples from participants who employed such a strategy in "Ernie" text (see Appendix 1). As most of the think-aloud comments were produced in L1 Japanese, the following examples were translated into English. The parts originally produced in English are underlined in italics. The comments related to thematic inference are in bold letters.

> *The interview* was long, and *Ernie thought he had done well*. He… *done well*…worked very hard and he thought he would be successful. (at Sentence 2) *He was sure that he would be soon employed as a security officer*. Hmm…*employed*…he thought he could get the job, *security officer*. Ah…**the theme of this story may be that working hard is important, right?** (at Sentence 3)…

> *… The next day he received a phone call* … he made a call? *from the factory manager saying he was not selected as a security guard position* he could not get the position? (at

Sentence 5) *Ernie was disappointed that* Ah… **acting recklessly without thinking will end in failure… it will cause trouble**. (at Sentence 6)

The learner who produced the comments above set a possible theme when she read parts of the passage and was able to verify and revise the first theme into a more appropriate one at the end of the passage. However, similar patterns were hardly found in the present study and most of the learners generated thematic inferences only at the end of the passages, or in the post-reading thematic inference task. This trend supported the results of Experiment 1 and Experiment 2 in that thematic inference was strategic processing facilitated by the task instructions rather than automatic processes.

Second, as for word and sentence analyses, the participants in the Control condition produced significantly more comments than in the Task condition. Moreover, the comments on Analysis were produced more in the Lower group than the Upper group. These results were consistent with the prediction derived from Experiments 1 and 2 in that EFL learners allocated more cognitive resources to lower-level linguistic processing and less resources to higher-level conceptual processing when they were not give any specific task instructions. Given that the text comprehension was higher in the Task condition than the Control condition, the amount of cognitive resources to word and sentence analyses paid by the learners in the Control condition was excessive; therefore, the task instructions helped learners to distribute such excessive resources to other levels of processing.

Third, the proportion of comments on backward and predictive inferences was not significantly different between the Control and Task conditions. According to the results of questionnaire, the learners paid attention to causal relationships and characters' goals in the narrative stories; therefore, it was predicted that they would generate more inferences, especially backward inferences, during reading in the Task condition than the Control condition.

Contrary to this prediction, the proportion of comments on inferences was relatively low in general, and there was no significant difference according to the task instructions. This small proportion of inferences can be attributed to the experimental passages used in this study. The causal structure of experimental passages was so simple that the learners could be satisfied with their comprehension of texts and not engage in strategic processes for coherence building. As a result, the necessity of inference generation could be low regardless of task instructions. In other words, learners were likely to employ lower standards of coherence (van den Broek et al., 1995).

Specifically, explicit signals such as discourse markers (e.g., *as a result*, *however*, *finally*) and argument and lexical overlap made easier for learners to construct coherent representations even without filling gap by generating inferences. For example, learners were likely to easily understand that the reason why Phil kept dating others and delayed proposing to his girl, based on a connective *so* included in the following sentence: "Phil was afraid of responsibility, *so* he kept dating others and delayed proposing to her." In this sentence, learners could understand the causal relationships between explicitly stated events without generatinf causal inferences based on their knowledge.

Regardless of such characteristics of experimental passages, the results of think-aloud data demonstrated that the participants in Upper group made more comments on inference than those in Lower group. The following think-aloud comments were the example of "Burt" passage (see Appendix 1) produced by a low-proficiency learner in the Task condition. Her recall production rate of this passage was 28%

> *One day, a large box accidentally*…*accident* and *-ly,* accidentally, *fell on him and*, he broke his shoulder. That's too bad. *Burt was in a pretty bad conditio*n. *"Pretty"* does not mean "beautiful" here, it means… *He had to spend several months at home*, he stayed his

home in order to recover. While at home, he studied *electronics* … Ah… and take courses by mail.

The following comments of "Burt" passage were produced by a higher-proficiency learner in the Task condition. The comments related to backward and predictive inference are in bold letters. His recall production rate of this passage was 61%.

One day, a large box accidentally fell on him **the box might be at a high place**. *and broke his shoulder*… He hurt his shoulder. (at Sentence 2) *Burt was in a pretty bad condition*. Yes, it should be natural. **Because a large and heavy box hit his shoulder**. (at Sentence 3) *He had to spend several months*, *several month*… Oh, **he must be a serious condition**. (at Sentence 4) *he started reading about electronics* Why? Suddenly? **In order to start a new job? Does he have too much time at home?** … *and decided to take courses by mail*. **Will he get a qualification for something?** (at Sentence 5)

In the protocols of lower-proficiency learner, although many comments on translations and analyses on words and sentences were produced, it seems that she could understand the passage and maintain the coherence of the story to some extent. In the protocols of the higher-proficiency learner, on the other hand, he compensated explicit information in the text by connecting events causally (e.g., the character started reading about electronics because he wants to start a new job) and anticipating the incoming events (e.g., the character will get a qualification for something). As a result, the allocation of cognitive resources to higher-level conceptual processing led to the construction of well-organized and robust representations of the passage (i.e., higher recall production rate).

Although clear interaction between L2 reading proficiency and task instructions on cognitive processes during reading was not found in the present study, the correlation analysis provided the possibility that the effects of task instructions can partially differ according to the learners' proficiency level. The results suggested that while less proficient learners were likely to allocate more cognitive resources to lower-level linguistic processing (i.e., analyzing words and sentences) regardless of the task instructions, the task instructions allowed proficient learners to allocate more cognitive resources to higher-level conceptual processing (i.e., backward and predictive inferences). In other words, high-proficiency learners were able to change their processing according to their reading goals, whereas it was difficult for low-proficiency learners to control their processing as skilled readers did. Similar observations were found in some previous L1 and L2 studies, suggesting that low-working memory readers, less skilled readers, and nonnative readers faced difficulty in flexible control of resource allocation in accordance with their reading goals given by the task instructions, compared to high-working memory, skilled, and native readers, respectively (e.g., Bohn-Gettler & Kendeou, 2014; Horiba, 2000; Linderholm & van den Broek, 2002).

3.3.5 Conclusion of Experiment 3

The purpose of Experiment 3 was to examine whether the task instructions aimed at thematic inference generation change (a) Japanese EFL learners' reading goals measured by a questionnaire and (b) Japanese EFL learners' cognitive processes measured by a think-aloud task. The main findings can be summarized into the following points.

First, the task instructions aimed at thematic inference generation changed Japanese EFL learners' reading goals (RQ3-1). When they were given the task instructions, their reading goal was shifted from understanding about characters in the story to building global coherence of

texts, understanding causal relationships between textual information, and understanding characters' intentions in the texts.

Second, the task instructions aimed at thematic inference generation partially changed Japanese EFL learners' cognitive processes during reading (RQ3-2). Although the effects of task instructions appeared only in one part of the reading process, the effects of task instructions appeared especially in lower-level processing. When they were not given task instructions, they allocated too many cognitive resources to translation and analyses of each word and sentence. The task instructions helped learners to distribute such excess resources to other levels of processing.

Third, the task instructions aimed at thematic inference facilitated Japanese EFL learners' text comprehension as in Experiment 2 (Hypothesis). The task instructions helped Japanese EFL learners pay attention to important elements (i.e., action and outcome), which contributes to construction of globally coherent and robust representations of narrative texts.

Finally, combining the results of the questionnaire and the think-aloud task, the following conclusion can be drawn. Both proficient and less proficient learners tried to employ different cognitive processes according to task instructions, which consequently leads to decrease of lower-level processing in Task conditions. However, as the negative correlations between L2 reading proficiency and the think-aloud comments on Analysis showed, lower-proficiency learners tend to require more cognitive resources for lower-level linguistic processing. As a result, only high-proficiency learners were able to change their processing according to their reading goals, whereas it was difficult for low-proficiency learners to control their processing as much as skilled readers.

Chapter 4

Study 2: Superordinate Inference Generation in Expository Reading

4.1 Experiment 4: Understanding Implicit Superordinate Propositions Through Inference Generation in Expository Reading

4.1.1 Purpose and Research Questions

The first purpose of Experiment 4 is to examine whether Japanese EFL learners can understand implicit superordinate propositions in expository texts through inference generation. To investigate readers' inference generation, the present study manipulated the explicitness of superordinate propositions in experimental passages as in Experiment 1. In the present study, the superordinate proposition is defined as the hierarchically highest statement that subsumes the relationship between a sequence of statements in the text. The present study used expository texts with a problem/solution structure as experimental materials; therefore, the superordinate proposition can be the statement that summarized the problem and the appropriate solutions for it (see Materials for details). The present study compared the conditions in which the participants read expository texts, including the explicit superordinate proposition, and in which they read expository texts without such a statement. The present study used an inference verification task in which the participants were required to judge whether the target statements could be understood or suggested from the passage they had read. When the learners inferred implicit superordinate statement, they would judge the implicit statement as highly appropriate as they did the explicit statement.

The second purpose of this experiment was to investigate whether the presence of superordinate propositions affect learners' mental representations of expository texts. Comparing the conditions stated above can show the differences of mental representation constructed by learners. Therefore, Experiment 4 sets two research questions as follows.

RQ4-1: Do Japanese EFL learners understand implicit superordinate propositions in expository texts by generating superordinate inferences?

RQ4-2: Do mental representations constructed by learners differ according to the explicitness of the superordinate propositions in the text?

Regarding RQ4-1, although there were few studies on inference generation in expository texts, some L1 studies provided important findings regarding the construction of the globally coherent representation of expository texts. For example, Ritchey (2011) demonstrated that superordinate inferences were generated spontaneously, rather than strategically, during expository texts. In contrast, some previous L1 studies suggested the difficulty of constructing superordinate propositions by integrating information beyond the paragraph level (e.g., Brown et al., 1983; Williams et al., 1984), and the ability to infer the superordinate propositions can be developed in college students. Ushiro and colleagues, who targeted Japanese EFL college students, also suggested that learners lack the ability to construct a superordinate proposition that embraces the ideas distributed across paragraphs. Based on these results, it can be predicted that Japanese EFL learners had difficulty generating superordinate inferences in expository reading.

Regarding RQ4-2, the present study investigated the effects of the explicitness of the superordinate propositions on the mental representation of expository texts. Ushiro and colleagues directly compared the texts with and without explicit macropropositions and investigated the differences in the construction of mental representations by learners using an importance-rating task. The results suggested that the absence of a macroproposition hindered connections with some of the information in the mental representations. Therefore, the present study can also predict that mental representations constructed by EFL learners differ according to the explicitness of the superordinate propositions in the text.

4.1.2 Method

4.1.2.1 Participants

Participants were 20 Japanese undergraduate students (9 female and 11 male) from the same university. They were all Japanese EFL learners with different majors (e.g., engineering, international studies, medical science, sociology, etc.) with intermediate or advanced levels of English proficiency. They had studied English for more than six years. The data of two participants who did not complete the given tasks were excluded from the analysis. Thus, the following analyses are based on the results of 18 participants.

4.1.2.2 Materials

Experimental passages

The experimental passages were two expository texts: "Africa's Great Green Wall" (the GGW text) and "Natural Solutions" (the NS text), adopted from the reading section of the Second Grade STEP test (STEP, 2011, 2012). These passages are presented in Appendix 5. Table 4.1 shows the number of words, sentences, and readability of each text. Both passages have a problem/solution structure[4] in which a problem was first raised, and then solutions to the problem were presented (Meyer & Freedle, 1984). For example, in the GGW text, the growth of the Sahara Desert was stated as a problem; the Great Green Wall project, where trees were planted to create a "wall," was stated as a solution. The NS text explained the problem of malaria spread by female mosquitoes, and then presented the solution of using fish to reduce mosquito populations.

For the experimental purpose, each experimental passage had two versions: (a) *explicit version*, in which the last sentence of the passage described the superordinate macroproposition,

[4] In the present study, passages with a problem/solution structure were selected as experimental materials because this type of passage was organized better than the other types (e.g., collection, description), and the writers are likely to convey their intended message as the solution to the problems.

and (b) *implicit version*, in which such a sentence was deleted from the passage. In the explicit version, the sentence "The Great Green Wall has many good points for stopping the desert's growth" was inserted in the GGW text, and "Using fish to control mosquitoes is an effective way of reducing malaria" was inserted in the NS text as the last sentence in each passage. These macropropositions in the passages were created through discussion between the author and two graduate students majoring in English education. In both passages, a problem and a solution stated in the passage are integrated into these macropropositions.

Table 4.1

Outline of the Experimental Passages in Experiment 4

	Explicit				Implicit			
Texts	Word	Sentence	FRE	FKGL	Word	Sentence	FRE	FKGL
GGW text	392	22	59.9	9.3	379	21	58.8	8.9
NS text	387	25	58.2	8.9	375	24	58.9	8.9

Note. FKR = Flesch Reading Ease; FKGL = Flesch-Kincaid Grade Level. These values were provided by Microsoft Word 2013's readability measurement tools.

Inference verification task

Target statements for the inference verification task were created for each experimental passage. Nine statements were prepared for each passage. The details were as follows: (a) a *consistent* statement, which is consistent with a macroproposition constructed from the overall passage; (b) an *inconsistent* statement, which is consistent with part of the passage, but which is inconsistent with the overall passage; (c) an *off-topic* statement, which is inconsistent with the main topic of the text; (d) three *explicit* statements, which described literal information explicitly mentioned in the passage; and (e) three *inappropriate* statements, which described

information not mentioned or suggested in the passage.

It should be noted that consistent statements were explicitly stated in the explicit version of each experimental passage while they were not presented in the implicit version. Therefore, if participants generated superordinate inferences in the implicit passage and then encoded the inferences as part of the text memory, consistent statements are likely to be judged "yes" and evaluated as highly appropriate. As in Experiment 1, all statements were presented to participants' L1 (Japanese) in order to avoid the effects of the participants' surface text memory regarding word forms and sentence structure on the verification task. The examples of target statements are shown in Table 4.2, and the target statements for the NS text are shown Appendix 6.

Before the experimental study, a norming study was conducted to verify the validity of the statements. The participants were 10 graduate and undergraduate students who majored in English education; none of them participated in the experimental study. The participants read two experimental passages with the implicit version and rated whether each statement could be understood from the passage on a 5-point scale.

Table 4.2

Target Statements (the GGW Text) in the Inference Verification Task in Experiment 4

(a) Consistent	Great Green Wall は、砂漠化の拡大を防ぐ方法として有効である。
	[The Great Green Wall has many good points for stopping the desert's growth.]
(b) Inconsistent	Great Green Wall は、環境に悪い影響をもたらす方法である。
	[The Great Green Wall has a bad influence on the environment.]
(c) Off-topic	森林の木を切ることは、かなり重労働な作業である。
	[Cutting trees is very hard work.]
(d) Explicit	Great Green Wall では、果物の木を植えることが有効である。
	[Planting fruit trees is effective for the Great Green Wall project.]
	Global Environment Facility (GEF) はプロジェクトに寄付をした。
	[The Global Environment Facility (GEF) donated to the project.]
	Great Green Wall は、動物たちの居住地にもなる。
	[The Great Green Wall will become a home for animals.]
(e) Inappropriate	Great Green Wall は、石で巨大な壁を作るプロジェクトである。
	[The Great Green Wall project created a wall of rocks.]
	Great Green Wall は、ヨーロッパの国々が行っているプロジェクトである。
	[Countries across Europe participated in the Great Green Wall project.]
	Great Green Wall は 2010 年代に始まったばかりだ。
	[The Great Green Wall only began in the 2010s.]

4.1.2.3 Procedure

Participants were tested individually in sessions that lasted about 30 min. Two passages were counterbalanced across two conditions (e.g., explicit and implicit). The experimental passages were presented in a random order to each participant. In the reading section, SuperLab

4.5 software (Cedrus, US) was installed on a computer, and the participants read passages using the response pad RB-730 (Cedrus, US). The experimental phase followed the instruction and practice phases. In the instruction phase, the procedure was explained in Japanese; and, in the practice phase, participants read an example passage to confirm the procedure.

The procedure was the same as in Experiment 1. Before the appearance of each passage, the signal *Ready?* appeared at the center of the screen and participants pushed the "yes" key to begin reading. They read each passage, sentence by sentence, in a self-paced fashion and pressed the "yes" key to signal that they had understood each sentence, so that they could not look at prior sentences. Participants were asked to carefully read each sentence in order to complete the post-reading verification task; however, they were not given specific reading goals for inference generation. When the last sentence of a text disappeared from the screen, the target sentences for the inference verification task appeared on the screen following the presence of "***" for 1,000 ms. The participants were required to decide if the verification statement could be appropriate, or suggested from the overall passage rather than part of the passage, as soon as possible. The participants were not allowed to return to the passages when they answered this task. Each verification statement was presented on the screen until the participants answered using the "yes" or "no" keys. After that, a 5-point scale from 1 (*not appropriate*) to 5 (*appropriate*) appeared on the screen, and the participants answered using the appropriate numeric keys. Nine statements for each passage were presented in random order, and the same procedure was repeated for two passages.

4.1.2.4 Scoring and Analysis

To examine whether the participants generate superordinate inference in expository passages, (a) the percentage of positive responses (i.e., "yes" responses) and (b) mean 5-point verification ratings were calculated and compared between two text conditions (i.e., explicit

and implicit conditions). If participants generated superordinate inferences in the implicit passage and then encoded the inferences as part of text memory, consistent statements were likely to be judged "yes" and evaluated as highly appropriate. As a result, the verification data of a consistent statement should not differ between the explicit and implicit conditions. In addition, to investigate the effects of explicitness of the macropropositions on the other text information, verification on consistent, inconsistent, and off-topic statements were compared between the two text conditions.

4.1.3 Results

4.1.3.1 Yes-Response Rates for the Verification Task

Regarding the consistent statements, the results indicated that 17 (95%) out of 18 participants answered yes in the Explicit condition, while 13 (72%) out of 18 participants answered yes in Implicit condition. To investigate the relationships between the response for Consistent statements and text condition, a Fisher's exact test was conducted. The results showed that the yes-response rate did not significantly differ between the text conditions, $p = .177$, $\varphi = .298$.

A second analysis was conducted on the response rates for Inconsistent and Off-topic statements, indicating the response for Inconsistent statements significantly related to the text condition [28% for Explicit vs. 61% for Implicit, χ^2 (1) = 4.05, $p = .044$, $\varphi = .335$]. On the other hand, the response for Off-topic statements does not relate significantly to the text condition [17% for Explicit vs. 22% for Implicit, χ^2 (1) = 0.18, $p = .674$, $\varphi = .070$].

4.1.3.2 Five-Point Rating Scores for the Verification Task

Table 4.3 illustrates the mean rating scores for the verification task. A 3 (Statement: Consistent, Inconsistent, Off-topic) × 2 (Text: Explicit, Implicit) two-way ANOVA was

conducted on the rating scores. The results showed that the main effect of the Statement, $F(2, 34) = 82.02$, $p < .001$, $\eta_p^2 = .678$, and the interaction between the two factors, $F(2, 34) = 12.50$, $p = .041$, $\eta_p^2 = .171$, were significant (see Table 4.4).

Table 4.3

Five-Point Scale Ratings for Statements in the Inference Verification Task in Experiment 4

	Consistent			Inconsistent			Off-topic		
	M	95% CI	*SD*	*M*	95% CI	*SD*	*M*	95% CI	*SD*
Explicit	4.72	[4.46, 4.99]	0.57	2.78	[2.06, 3.50]	1.56	2.22	[1.73, 2.71]	1.06
Implicit	4.28	[3.76, 4.80]	1.13	3.28	[2.61, 3.95]	1.45	2.33	[1.83, 2.83]	1.08

Table 4.4

Summary Table for Two-Way ANOVA of the Effects of the Statement and Text on 5-Point Ratings in Experiment 4

Source	*SS*	*df*	*MS*	*F*	*p*	η_p^2
Text	0.33	1	0.33	0.23	.638	.013
Error (Text)	24.67	17	1.45			
Statement	82.02	2	41.01	35.77	.000	.678
Error (Statement)	38.98	34	1.15			
Text × Statement	12.50	2	6.25	3.51	.041	.171
Error (Text × Statement)	60.50	34	1.78			

Follow-up ANOVAs with Bonferroni correction were conducted on rating scores for each text condition to examine whether the rating scores were different according to the explicitness of the superordinate proposition. The results indicated that the main effect of Statement was

significant in both Explicit and Implicit conditions. The results indicated that the rating score for Consistent was significantly higher than Inconsistent and Off-topic in the Explicit condition (*p*s < .001). In the Implicit condition, the rating scores for Consistent and Inconsistent were not significantly different, but they were significantly higher than that of the Off-topic statements (p = .003) (see Figure 4.1). Moreover, the follow-up analysis also indicated that rating scores for Consistent were likely to be higher in the Explicit condition compared to the Implicit condition (p = .055).

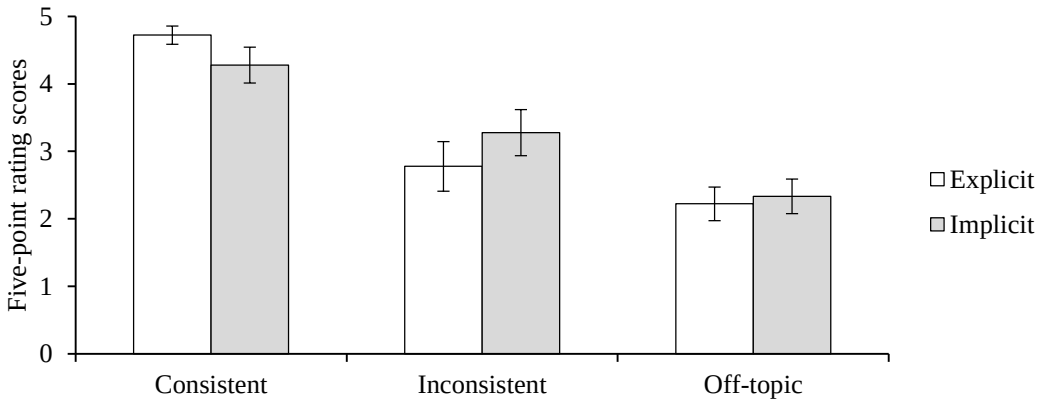

Figure 4.1. Five-point rating scores by statement and text condition in Experiment 4.

4.1.4 Discussion

Do Japanese EFL learners understand the implicit superordinate propositions in expository texts by generating superordinate inferences? (RQ4-1)

The results of the verification task showed that the yes-response rate and 5-point ratings for a Consistent statement did not differ between Explicit and Implicit texts. The descriptive statistics also indicated that although the ratings in Explicit texts were slightly higher than those in Implicit texts, most participants judged the Consistent statements as appropriate or suggested them from the overall passage regardless of their explicitness. Therefore, the results suggested that Japanese EFL learners are likely to comprehend the implicit superordinate proposition

through inference generation. In fact, these Consistent statements are rated as highly as Explicit statements[5] (literal information explicitly mentioned in the passage; yes-response rate: *M* = 87.96%, *SD* = 12.53; 5-point ratings: *M* = 4.21, *SD* = 0.41). In other words, the participants encoded both the explicit and implicit superordinate propositions into their mental representations as strongly as the explicitly stated information in the text.

The current study results are partly consistent with Ritchey (2011), who investigated the generation of superordinate inferences among L1 readers. Ritchey showed that readers generated superordinate inferences regardless of different reading goals: summarization task—which is used to encourage superordinate inference generation—and verification task—which is used to process details. The results showed a similar pattern for reading time on target sentences, suggesting that generation of superordinate inferences was mandatory for reading comprehension, rather than the occurrence of optional and strategic processing under specific conditions. The present study did not intend for the participants to engage in the specific goal of constructing global coherence of texts, but they were told to answer the verification task after reading each passage. Nevertheless, they judged the implicit superordinate proposition as highly appropriate as they did the explicit one. Therefore, the present study results can support the notion that superordinate inferences are generated spontaneously rather than strategically without any specific tasks.

The present study results were not consistent with Ushiro and colleagues (2008), who revealed that EFL learners had difficulty constructing implicit macropropositions based on the information across paragraphs. This inconsistency can be attributed to the methodology adopted in these studies. In Ushiro and colleagues, the construction of an implicit macroproposition was measured by whether the participants included implicit macropropositions in their summary

[5] As each text has three Explicit statements and one Consistent statement, these statements were not directly compared.

protocols. The present study required the participants to only judge the presented statements rather than create implicit superordinate propositions by themselves, which might lead to a higher yes-response rate and 5-point verification scores. In other words, they might not activate superordinate inferences until the consistent statement was presented as the target statement in the verification task.

Do mental representations constructed by learners differ according to the explicitness of the superordinate propositions in the text? (RQ4-2)

The results of the verification task indicated that the verification pattern was partly different between the Explicit and Implicit texts. Specifically, the rating score for Consistent was significantly higher than Inconsistent and Off-topic in the Explicit text, whereas the rating scores for Consistent and Inconsistent were not significantly different but were significantly higher than that for Off-topic in the Implicit text. When the superordinate macroproposition, which subsumes the relationship between a sequence of statements in the text, was not explicitly stated in the expository text, learners were likely to regard the Inconsistent statements as highly appropriate as they would Consistent statements. These results suggested that learners could not suppress the activation of information, which is consistent with the local comprehension of the text but inconsistent with the overall passage. For example, during reading the GGW text, the readers understand the information "The Great Green Wall has a bad influence on environment" as the Inconsistent statements are presented in the verification task based on the third paragraph. The third paragraph of the GGW text is as follows:

> "A number of NGOs representing local communities, however, are concerned about the plan's possible effects. They are especially worried about the idea of planting a large number of trees. They say this might mean introducing new kinds of trees from other

areas that would damage local ecosystems and use up valuable farmland."

However, if the readers constructed globally coherent representation of the text by integrating information between the third paragraph and previous/following paragraphs, the memory of the information that is inconsistent with the overall text representation (e.g., the shortcomings of the Great Green Wall) would become weak. The results of the verification task demonstrated that when the participants read the Explicit text, they regarded the Inconsistent statement as less appropriate compared to the Consistent statement. When they read the Implicit text, they regarded the Inconsistent statement as appropriate as the Consistent statement. These results suggested that the absence of a superordinate proposition could hinder the construction of globally coherent representations of the text. The similar finding was also derived from Ushiro and colleagues (2008) in that when the superordinate information was not explicitly stated in the expository text, the learners were not able to appropriately connect some pieces of information in the text.

Such difficulty in suppressing unimportant and irrelevant information when they read the text without superordinate information can be explained from the construction-integration model (Kintsch & van Dijk, 1978) and landscape model (van den Broek et al., 1996). These models assumed that readers activated text information and background knowledge to construct mental representations first; then, the irrelevant information was deactivated and disappeared from their representations. These activation or deactivation patterns were dependent on the various sources (see section 2.1.1). In the case of the present study, the difference between the Explicit and Implicit condition was only in that presence or absence of a superordinate proposition. When the superordinate proposition was explicit, the relevant information (e.g., good points of the Great Green Wall) was likely to activate while irrelevant information (e.g., shortcomings of the Great Green Wall) was deactivated. When the superordinate proposition

was not explicitly stated, the relevant information was not likely to activate enough, and irrelevant information was not deactivated. As a result of such activation and deactivation, the absence of the superordinate proposition affected the mental representations constructed from the texts.

In sum, the results suggested that mental representations constructed by learners can be affected by the explicitness of the superordinate propositions in the text. When the superordinate proposition was explicitly stated, the mental representation of the text was constructed around the superordinate proposition. Conversely, when the superordinate proposition was implicit, the unimportant information was still activated; consequently, it led to less coherent mental representations of the text. Therefore, Experiment 4 confirmed the significance of understanding superordinate propositions for building globally coherent representations of expository texts.

4.1.5 Conclusion of Experiment 4

Experiment 4 aimed to examine whether Japanese EFL learners can understand the implicit superordinate propositions in expository texts through inference generation, and whether the presence of the superordinate propositions affect the learners' mental representations of expository texts. The findings of this experiment can be summarized as follows. First, EFL learners were able to understand implicit superordinate propositions in expository texts by generating superordinate inferences. Second, mental representations constructed by EFL learners can be affected by the explicitness of the superordinate propositions in the text.

This experiment is the first step of investigating superordinate inference generation in expository reading and there are some limitations to solve. First, the present study required participants to only judge the presented statements rather than answer the superordinate

propositions created on their own. Therefore, the participants might judge implicit statements as relatively high. Therefore, the following studies need to replicate the generation of superordinate inference by adopting other methodology. Experiment 5 directly measured inference generation using a superordinate inference task.

Second, whether the generation of superordinate inference was mandatory for reading comprehension should be examined further. Although the present study suggested that generation of superordinate inferences can be mandatory for reading comprehension because they judged the implicit superordinate propositions as highly appropriate as the explicit ones without specific instructions. However, as reviewed in section 2.3, readers' inference generation and text comprehension are strongly influenced by task instructions. Therefore, the following studies investigated whether the task instructions facilitated superordinate inference and text comprehension. If the generation of superordinate inference was an automatic process for reading comprehension, the inference generation would not be influenced by the task instructions.

Third, the effects of generating superordinate inference on text comprehension are still unclear. The results of Experiment 4 suggested that the explicitness of the superordinate propositions can influence on the mental representations constructed by learners. However, as the present study used only the inference verification task, the following studies need to adopt other methodologies to measure text comprehension in order to investigate whether the generation of superordinate inferences facilitated text comprehension.

To address these issues, Experiment 5 manipulated the task instructions in order to investigate whether the generation of superordinate inferences was a mandatory or strategic process in reading comprehension. Specifically, participants were randomly divided into two task conditions as follows. Half of the participants were told to read expository texts while thinking about what the writer intended to convey through the overall text (Task condition); the

other half were told to read the passages in order to answer comprehension questions after reading. Moreover, to directly investigate the generation of superordinate inference, Experiment 5 required the participants to answer the superordinate inference task in which the participants answered the message conveyed through the overall expository texts in one sentence in Japanese after reading the texts. Finally, learners' text comprehension was measured by a written recall task to investigate the effects of generating superordinate inference on text comprehension.

4.2 Experiment 5: Effects of Task Instructions on Superordinate Inference and Text Comprehension

4.2.1 Purpose and Research Questions

Experiment 4 demonstrated that EFL learners were able to understand the implicit superordinate propositions in expository texts by generating superordinate inference. Moreover, the explicitness of the superordinate propositions in the text can affect the mental representations constructed by EFL learners. Based on these findings, Experiment 5 aimed to (a) replicate the generation of superordinate inferences and (b) examine the effects of task instructions on inference generation and text comprehension.

Experiment 4 suggested that Japanese EFL learners were able to generate superordinate inferences without any strategic instructions. However, as Experiment 4 used an inference verification task in which participants were simply required to judge the appropriateness of the presented statements, superordinate inference generation should be replicated again in the present study. Therefore, the present study directly measured inference generation by using a superordinate inference task. In this task, the participants answered the message conveyed through the overall expository texts in one sentence in Japanese after reading the texts.

Moreover, the present study aimed to investigate whether the task instructions aimed at superordinate inference affect inference generation and text comprehension. Regarding inference generation, Experiment 4 suggested that generation of superordinate inferences can be mandatory for reading comprehension, even though they were not told to do so (i.e., without any specific task instructions). However, as stated above, learners may have difficulty constructing superordinate propositions by themselves. In that case, the task instructions would facilitate superordinate inference generation. In order to examine this possibility, Experiment 5 manipulated the task instructions as in Experiment 2. Specifically, the present study provided a reading goal for constructing superordinate propositions to learners; they were asked to read

passages in order to understand what the writer intended to convey through the overall text.

Regarding text comprehension, the results of Experiment 4 suggested that understanding superordinate propositions can influence the mental representations constructed by learners. In order to examine this point further, Experiment 5 adopted a written recall task as a measurement of text comprehension.

RQ5-1: Do task instructions facilitate EFL learners' superordinate inference generation in expository reading?

RQ5-2: Do task instructions aimed at superordinate inference affect Japanese EFL learners' text comprehension?

As reviewed in section 2.3, given a specific goal by task instructions, readers engage in strategic processing according to the reading goals that affect the processes and products of reading (e.g., Geiger & Millis, 2004; Kaakinen & Hyönä, 2005; Kaakinen et al., 2002; Linderholm & van den Broek, 2002; Magliano et al., 1999; van den Broek et al., 2001). Therefore, when the learners were given task instructions, which induced them to construct superordinate propositions subsuming overall text information, they would alter their reading processes to achieve their goals, which lead to facilitation of superordinate inference generation and text comprehension.

Conversely, L2 and EFL learners have greater difficulty controlling their text comprehension according to the task instructions compared to L1 readers (e.g., Horiba, 2000, 2013; Yoshida, 2012). Therefore, task instructions used in the present study would have small, or even no, influence on EFL learners' inference generation and text comprehension or, possibly, only high proficiency learners benefit from the task instructions.

4.2.2. Method

4.2.2.1 Participants

Participants were 68 Japanese undergraduate and graduate students (30 female, 38 male) who majored in a variety of fields, such as social studies, engineering, biology, and medical science. Data from one participant who could not appropriately complete the task was excluded; thus, data from 67 participants were available for full analysis. All participants had studied English for more than six years in Japan, and their English proficiency levels ranged from intermediate to advanced.

4.2.2.2 Materials

L2 reading proficiency test

In order to assess the participants' English reading ability, a 24-item reading proficiency test was prepared. The passages and questions were the same as in Experiments 2 and 3.

Experimental passages

Two expository texts used in Experiment 4 were also used as the experimental passages in this study. However, the present study did not aim to investigate the effects of the explicitness of the superordinate propositions in the text; therefore, this study only adopted the Implicit version of each text.

4.2.2.3 Procedure

Task instructions

To investigate the effects of task instructions on superordinate inference generation and text comprehension, the participants were assigned to either of the following two reading conditions. In the Task condition, before reading the texts, participants were asked to think

about what the writer intended to convey through the overall text. They were told that they would answer a superordinate inference task after reading. In the Control condition, participants were told to read the passages in order to answer comprehension questions after reading. This instruction would not alter any strategic processing for generating superordinate inference because they were not given any information about the superordinate inference task. In both conditions, the participants were not informed of a later recall task.

Experimental sessions

The participants were tested in a group setting (4 to 10 participants per group). First, participants completed the reading proficiency test in 25 minutes. Then, participants were randomly assigned to one of the two reading conditions: reading for the superordinate inference task (the Task condition) or reading for text comprehension (the Control condition).

Although the pre-reading instructions were different according to the task condition, the participants in both conditions completed the same procedure. First, they read each passage silently for 4 minutes. The presentation orders of the two passages were randomized across the participants. The time limit for reading was determined based on the pilot study with a small group of participants, and the participants of Experiment 5 reported they finished reading each passage in the time limit. Immediately after reading the two passages, participants began a written recall task in which they wrote down, in Japanese, all of what they remembered. Participants were asked to complete the written recall task in ten minutes, but they were given additional time if necessary in order to collect as much data as possible. Finally, participants completed a superordinate inference task for the two passages in 5 minutes. This task required participants to write down, in Japanese, the message conveyed through the overall expository texts in one sentence in Japanese after reading the texts.

4.2.2.4 Scoring and Analysis

Superordinate inference task

Learners who could appropriately describe the superordinate proposition (i.e., "The Great Green Wall has many good points for stopping the desert's growth" for the GGW text, and "Using fish to control mosquitoes is an effective way of reducing malaria" for the NS text) were categorized as "successful," while those who could not were categorized as "failed." In addition, failed learners' answers were further analyzed qualitatively; these answers were categorized into (a) broad and (b) narrow. A broad statement indicates that readers provided too broad or general statement such as, "Solving environmental problems is really difficult (the GGW text)." Failed readers could also write a sentence that was too narrow in focus that described only part of the text, such as, "Planting fruit trees offers a source of income, which helps the local people (the GGW text)." Based on the criteria, 30% of the data were randomly selected and scored by two raters, with an agreement rate of 82.50%. Disagreements were resolved through discussion, and the researcher scored the remaining data alone.

Written recall task

First, experimental passages were divided into a set of idea units (IUs) based on Ikeno's criteria (1996) as in Experiment 2. Two raters conducted this division, and the agreement between them was 99.34%. Disagreements were resolved through discussion. The total number of IUs was 73 for the GGW text and 81 for the NS text. 30% of recall data was randomly selected and scored by two raters, separately. If two-thirds of the information in the IU was reproduced in a participant's recall protocol, one point was given to that IU. The agreement between the two raters was 95.19%. Disagreements were resolved through discussion, and the researcher scored the remaining data alone.

4.2.3 Results

4.2.3.1 L2 Reading Proficiency Test

The reliability of the reading proficiency test was acceptable (Cronbach's α = .78); therefore, participants were divided into two proficiency groups (Upper, Lower) according to a median split of the test scores. The number of participants and the mean scores are shown in Table 4.5. In order to confirm whether the proficiency level of the two reading conditions was homogeneous, a 2 (Proficiency: Upper, Lower) × 2 (Condition: Control, Task) ANOVA was conducted on proficiency test scores (see Table 4.6). Results indicated a significant main effect of Proficiency, F (1, 63) = 198.59, $p < .001$, $\eta_p^2 = .759$, whereas the main effect of Condition, $F(1, 63) = 0.07$, $p = .799$, $\eta_p^2 = .001$, and the interaction between these two factors, $F(1, 63) = 1.49$, $p = .226$, $\eta_p^2 = .023$, were not statistically significant. These results confirmed that the L2 reading proficiency level was not significantly different between the two reading conditions.

Table 4.5

Descriptive Statistics of L2 Reading Proficiency Test in Experiment 5

	Proficiency	*n*	*M*	95% CI	*SD*
Control	Upper	18	20.44	[19.65, 21.24]	1.72
	Lower	15	14.07	[12.88, 15.25]	2.34
	Total	33	17.55	[16.25, 18.84]	3.79
Task	Upper	17	21.18	[20.37, 21.99]	1.70
	Lower	17	13.59	[12.50, 14.68]	2.29
	Total	34	17.38	[15.93, 18.84]	4.34

Note. Maximum possible score is 24.

Table 4.6

Summary Table for Two-Way ANOVA of the Effects of Condition and Proficiency on the Score of L2 Reading Proficiency Test in Experiment 5

Source	*SS*	*df*	*MS*	*F*	*p*	η_p^2
Condition	0.27	1	0.27	0.07	.799	.001
Proficiency	813.15	1	813.15	198.59	.000	.759
Condition × Proficiency	6.11	1	6.11	1.49	.226	.023
Error	257.97	63	4.09			
Total	21512.00	67				

4.2.3.2 Superordinate Inference Task

Effects of task instructions on superordinate inference

First, participants' answers were categorized into (a) successful and (b) failed in order to examine how many learners in each reading condition succeeded within superordinate inference generation. Cross-tabulations for each passage were used to examine the relationship between success (i.e., successful and failed) and reading conditions using the data shown in Table 4.7.

Although the number of successful readers in the Task condition seemed to be larger than that in the Control condition, two-way chi-square tests showed that the number of learners who succeeded with inference generation did not significantly differ by reading condition for both texts.

Table 4.7

Number of Successful and Failed Participants in the Superordinate Inference Task by Reading Condition in Experiment 5

	Control (n = 33)		Task (n = 34)				
	Successful	Failed	Successful	Failed	χ^2	p	ϕ
GGW text	17	16	23	11	1.81	.178	.164
NS text	12	21	19	15	2.57	.109	.196

Effects of L2 reading proficiency on superordinate inference

Second, cross-tabulations were used to examine the relationship between success and L2 reading proficiency using data shown in Table 4.8. The table indicates that higher proficiency learners tended to succeed with superordinate inference, whereas lower proficiency learners tended to fail. Two-way chi-square tests showed that superordinate inference for both texts significantly differed by L2 reading proficiency: Readers in the higher proficiency group were more likely to succeed with the superordinate inference task than those in the lower group.

Table 4.8

Number of Successful and Failed Participants in the Superordinate Inference Task by L2 Reading Proficiency in Experiment 5

	Upper (n = 35)		Lower (n = 32)				
	Successful	Failed	Successful	Failed	χ^2	p	ϕ
GGW text	25	10	15	17	4.19	.041	.250
NS text	21	14	10	22	5.56	.018	.288

Qualitative analysis of inaccurate statements

In addition to quantitative analyses of superordinate inference task, this section discussed whether superordinate inference differed qualitatively according to task conditions. Readers' failed answers were categorized into (a) broad and (b) narrow. Table 4.9 shows the answer pattern by each reading condition, and Table 4.10 presents example answers.

Table 4.9

Answer Patterns of Failed Participants in Experiment 5

	Control		Task	
	Broad	Narrow	Broad	Narrow
GGW text	6	10	2	9
NS text	9	12	3	12

Most failed learners reported narrow statements, which describe only part of the text rather than the overall text. This tendency was found in both the Control and Task conditions, which indicated that qualitative differences in superordinate inference were not detected between two reading conditions: Readers in both conditions tended to construct narrow statements.

Examples (b) and (c) in Table 4.10 indicate that these learners seem to be able to understand passage topics but are not able to identify problems and solutions in the texts. Meanwhile, examples (d) and (e) explicitly stated sentences based on the experimental passage (e.g., a topic sentence for each paragraph) rather than the superordinate proposition of the whole text. Such answers indicated that these learners have the ability to construct a macroproposition from a single paragraph but have difficulty integrating information from the overall text.

Table 4.10

Example Answers for the Superordinate Inference Task in Experiment 5

	the GGW text	the NS text
Successful	(a) Creating the Great Green Wall by planting trees is an effective method for stopping the desert's growth.	(a) Biocontrol, which uses nature to control nature, is a good solution to malaria.
Failed: Broad	(b) It is necessary but difficult to stop the desert's growth. (c) People all over the world should have an interest in environmental preservation.	(b) Controlling nature is good for people. (c) Solving environmental problems are really difficult.
Failed: Narrow	(d) The Great Green Wall provides a good environment for plants and animals. (e) Planting fruit trees offers a source of income, which helps the local people.	(d) Biocontrol is especially effective in developing countries where malaria is still common. (e) A researcher found a new type of fish that can survive even without water.

Note. The original answers were in Japanese and translated into English by the author.

4.2.3.3 Written Recall Task

The mean production rates for the written recall task were shown in Table 4.11. The production rates were converted into percentages (i.e., the number of recalled IUs out of the total) because passage length was different between two texts. A 2 (Text: GGW, NS) × 2 (Condition: Control, Task) × 2 (Proficiency: Upper, Lower) three-way ANOVA was conducted on the mean production rates with Text as a within-subjects factor, and Condition and

Proficiency as between-subjects factors (see Table 4.12). Results showed that only the main effect of reading proficiency was significant, $F(1, 63) = 29.67$, $p < .001$, $\eta_p^2 = .320$, whereas all other main effects and interactions were not significant (ps > .10). These results indicate that Upper group recalled better than Lower group regardless of the task instructions and texts (see Figures 4.2 and 4.3).

Table 4.11

Mean Production Rates on the Recall Task in Experiment 5

			GGW text			NS text		
Condition	Proficiency	*n*	*M*	95% CI	*SD*	*M*	95% CI	*SD*
Control	Upper	18	30.03	[25.40, 34.67]	10.04	30.45	[26.79, 34.11]	7.93
	Lower	15	19.42	[14.14, 24.70]	10.43	21.52	[16.09, 26.95]	10.73
	Total	33	25.21	[21.32, 29.10]	11.40	26.39	[22.91, 29.87]	10.21
Task	Upper	17	30.78	[25.72, 35.83]	10.63	31.65	[27.66, 35.63]	8.39
	Lower	17	17.90	[13.98, 21.83]	8.26	19.43	[15.24, 21.81]	8.82
	Total	34	24.34	[20.50, 28.18]	11.43	25.54	[22.05, 27.48]	10.50

Table 4.12

Summary Table for Three-Way ANOVA of the Effects of Text, Proficiency, and Condition on the Recall Production Rates in Experiment 5

Source	*SS*	*df*	*MS*	*F*	*p*	η_p^2
		Within-participants				
Text	50.41	1	50.41	1.34	.251	.021
Text × Condition	0.03	1	0.03	0.00	.978	.000
Text × Proficiency	11.42	1	11.42	0.30	.584	.005
Text × Condition × Proficiency	2.18	1	2.18	0.06	.811	.001
Error (Text)	2369.79	63	37.62			
		Between-participants				
Condition	5.78	1	5.78	0.04	.840	.001
Proficiency	4151.78	1	4151.78	29.67	.000	.320
Condition × Proficiency	63.95	1	63.95	0.46	.502	.007
Error	8816.74	63	139.95			

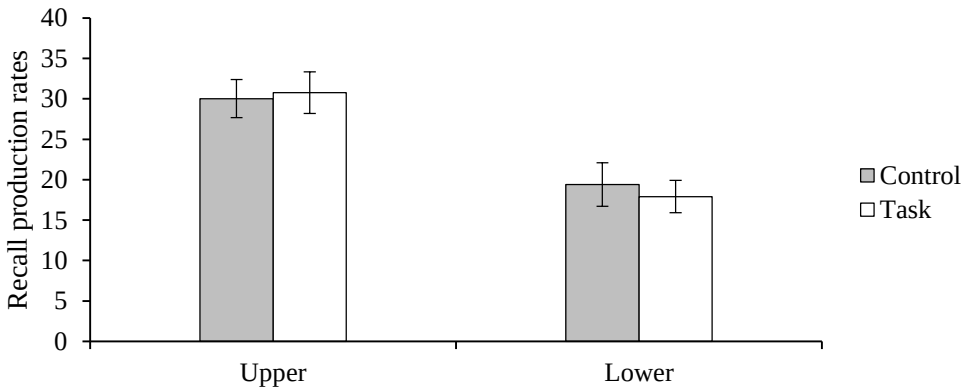

Figure 4.2. Total recall production rates (%) for the GGW text (Condition × Proficiency) in Experiment 5. Error bars represent standard errors.

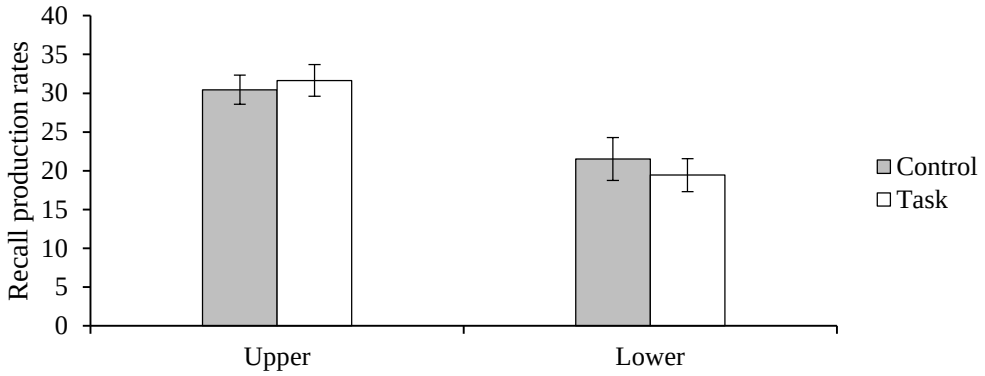

Figure 4.3. Total recall production rates (%) for the NS text (Condition × Proficiency) in Experiment 5. Error bars represent standard errors.

4.2.4 Discussion

The effects of task instructions on superordinate inference generation (RQ5-1)

The results of the superordinate inference task demonstrated that the task instructions did not relate to the success in superordinate inference generation. Moreover, the qualitative analysis showed a similar pattern of superordinate inference between the two reading conditions. Learners who failed with superordinate inference tended to construct narrow statements that corresponded to the main idea of each paragraph rather than the message of the overall text.

For example, one answer "The Great Green Wall provides a good environment for plants and animals" was explicitly written information in the second paragraph of the GGW text, but this was not the central message conveyed through the overall text. Similarly, an example answer "Planting fruit trees offers a source of income, which helps the local people" was explicitly mentioned in the fourth paragraph, but it did not subsume the overall contents of the passage. Thus, learners who failed in superordinate inference were likely to answer the main idea of one paragraph as the writer's message conveyed through the overall text. In other words, absence of superordinate propositions can prompt learners to construct a mental representation around the local main idea (i.e., the main idea of each paragraph) rather than around the

hierarchically highest macroproposition. Consequently, the learners were unable to grasp an overall picture (i.e., problem/solution structure) of the expository texts.

A similar result was found by Ushiro and colleagues (2008), who showed that learners were able to construct macropropositions in each paragraph whereas they had difficulty integrating and constructing information across paragraphs. Therefore, the task instructions provided in the present study (i.e., to think about what the writer intended to convey through the overall text, and answer it in a post-reading superordinate inference task) was insufficient in facilitating integrating information across paragraphs, which lead to failure in constructing globally coherent representation of expository texts. Based on these results, EFL learners' superordinate inference generation can be facilitated when they are instructed to integrate information distributed across paragraphs. Therefore, the effects of task instructions on superordinate inference generation should be further examined in the next experiment.

Conversely, success in superordinate inference generation was significantly related to learners' English reading proficiency. Learners with higher proficiency succeeded at the superordinate inference task more compared to learners with lower proficiency. Previous L1 research has also demonstrated that adult readers constructed appropriate macrostructures of the text better than children (Brown et al., 1983; Williams et al., 1984). Moreover, although L2 reading research comprehensively examined various types of inferences rather than directly focused on superordinate inferences, some studies revealed that L2 reading proficiency was related to inference generation (Horiba, 1996; Muramoto, 2000; Yoshida, 2003). The present study results supported these studies and demonstrated one of the different characteristics between good and poor EFL readers.

The effects of task instructions on text comprehension (RQ5-2)

Experiment 5 adopted a written recall task as a measurement of text comprehension in

order to further investigate the possibility that understanding superordinate propositions can influence the mental representations constructed by learners. The results of the written recall task showed the main effect of L2 reading proficiency, whereas neither the main effect of reading condition nor an interaction between these two factors was significant. It was demonstrated that learners with higher L2 reading proficiency recalled the expository texts better compared to those with lower proficiency regardless of the task instructions. The reason why task instructions did not affect text comprehension can be explained by the following two points.

First, the learners had difficulty controlling their reading comprehension according to the reading goals given by the task instructions. As reviewed in section 2.3, L1 readers were able to adjust their cognitive processes and strategies in accordance with their reading goals, which affected products of text comprehension measured by the recall task (e.g., Linderholm & van den Broek, 2002; Magliano et al., 1999; van den Broek et al, 2001). However, Horiba (2000), which examined the task instructions on reading processes and text comprehension by direct comparison between L1 and L2 reading, revealed that L2 readers had more difficulty controlling their reading processes according to the task instructions compared to L1 readers. In the present study, such difficulty for L2 readers might lead to similar recall production rates between the Task and Control conditions. However, as the task instructions did not influence superordinate inference as well as text comprehension, the following possibility seems more plausible.

The second possibility is about the characteristics of the task instructions. As reviewed in section 2.3.2, whether L2 text comprehension is influenced by the task instructions is likely to depend on the specificity of the task instructions. Specifically, L2 text comprehension was not influenced by the task instructions (e.g., Horiba, 2013; Yoshida, 2012), especially when learners were given instructions that did not require them to alter specific types of processing

during reading (e.g., task instructions aimed at specific type of inference). As in Horiba (2000) and Nahatame (2014), when learners were given task instructions that required learners to alter specific types of processing during reading (e.g., reading for coherence, reading for predictive inference generation), they were able to strategically control their inference generation according to the given instructions, which lead to positive influence on text comprehension. Although the present study aimed to induce learners to generate superordinate inferences, the task instructions might be somewhat ambiguous and less strategic for learners; thus, both inference generation and text comprehension were not influenced by the task instructions. Therefore, to induce learners to generate superordinate inference, giving a more specific and strategic goal might be more effective. For example, as the results of the superordinate inference task suggested, the task instructions in which learners are instructed to integrate information distributed across paragraphs can alter the strategic processing of learners.

4.2.5 Conclusion of Experiment 5

The purpose of Experiment 5 was to replicate the generation of superordinate inferences, and examine the effects of task instructions on inference generation and text comprehension. The present study adopted the superordinate inference task as a measurement of inference generation, and the written recall task as a text comprehension. The main findings of the present study can be summarized into the following three points: (a) the superordinate inference was not routinely and spontaneously constructed but it was affected by the learners' L2 reading proficiency, and (b) the task instructions used in the present study did not affect superordinate inference generation and text comprehension.

First, the superordinate inference task demonstrated that learners had difficulty constructing superordinate propositions that subsume the relationship between a sequence of statements in the text. In Experiment 4, most participants judged implicit superordinate

propositions as appropriate, suggesting that they can generate superordinate inference without any task instructions. However, the reason why they were able to generate superordinate inference can be attributed to the inference verification task used in Experiment 4, which required participants to only judge the presented statements. Therefore, Experiment 5 aimed to replicate the inference generation using the superordinate inference task.

The results of the superordinate inference task demonstrated that around half of the participants failed in superordinate inference generation when not told to do so (i.e., the Control condition). It was suggested that superordinate inference was not likely to be mandatory for reading comprehension and that it can occur under a particular condition. Although the present study did not find the relationship between task instructions and superordinate inference generation, L2 reading proficiency was significantly related to superordinate inference generation. It can be concluded that superordinate inference is not a routine and spontaneous process that necessarily occurs under various circumstances; rather, it characterizes the ability of skilled L2 readers. The following experiment needs to further examine effective task instructions, which would facilitate learners' superordinate inference generation.

Second, the task instructions used in the present study (i.e., read the texts while thinking about what the writer intended to convey through the overall text) did not affect either superordinate inference generation or text comprehension. The results of superordinate inference task demonstrated that learners were likely to answer the main idea of one paragraph as the writer's message conveyed through the overall text. In other words, they had difficulty integrating information distributed across paragraphs. The task instructions given in the present study might be less strategic for supplementing this difficulty by altering learners' strategic processing. Therefore, the effects of task instructions should be further investigated in the next experiment. When the task instructions required learners to alter specific types of processing during reading, text comprehension and inference generation would be facilitated.

In sum, Experiment 5 demonstrated that it was difficult for EFL readers, especially for less skilled readers, to construct implicit superordinate propositions by inference generation. Moreover, the task instructions in which learners were only told to think about the writer's message were insufficient in altering the learners' reading goals. Learners were likely to lack the ability to integrate information across paragraphs; therefore, task instructions that compensate for the deficit should be considered. To construct a global coherence of the text, readers must link currently processing elements and much earlier elements no longer kept in (short-term) working memory and/or to relevant world knowledge (Graesser et al., 1994). The information in the previous paragraphs is not activated in the readers' working memory; therefore, to integrate information across paragraphs, readers must retrieve the information from their long-term memory. Therefore, Experiment 6 manipulated the task instructions that aimed to facilitate such processes, and reexamined the task instructions on superordinate inference generation and text comprehension.

In addition, as reviewed in section 2.3.2, the effects of task instructions on L2 reading comprehension are more complex and unstable compared to those of L2 reading comprehension; therefore, not only the products of comprehension (i.e., the performance of a recall task) but also the products of comprehension should be examined. Experiment 6 examined EFL learners' reading processes by adopting a think-aloud task as in Experiment 3. It will provide more detailed data to assess learners' control of comprehension while reading expository texts.

4.3 Experiment 6: Effects of Integration Task on Processes and Products of Expository Comprehension

4.3.1 Purpose and Research Questions

The aim of Experiment 6 was to examine the following three points: whether a particular reading task that induced EFL learners to integrate information were distributed among paragraphs on (a) superordinate inference, (b) on-line processes during reading, and (c) text comprehension after reading.

Combining the results of Experiments 4 and 5, it was suggested that Japanese EFL learners had difficulty constructing implicit superordinate propositions by inference generation. Experiment 5 demonstrated that such difficulty by learners can be attributed to the lack of ability to integrate information across paragraphs. Therefore, Experiment 6 manipulated the task instructions, which aimed to facilitate such processes, and reexamined the effectiveness of the reading task. Specifically, the present study gave the integration task in which the participants were told to answer what the writer intended to convey, every time they finished reading each paragraph. Thus, when they finished reading the first paragraph, they answered the writer's message conveyed through the first paragraph; then they finished reading the second paragraph and answered the writer's message by integrating information across the first and second paragraphs. The same procedure followed when they finished reading the third and fourth paragraphs. The detailed task instructions and procedure are explained in the Methods.

As in previous studies and Experiment 3, a think-aloud method was adopted in this experiment to gain insight into the cognitive processes and strategies learners used during reading, whereas the written recall task was adopted to measure the products of reading comprehension. The three research questions (RQs) are summarized as follows:

RQ6-1: Does the integration task facilitate superordinate inference generation?

RQ6-2: Does the integration task affect EFL learners' processes of text comprehension?

RQ6-3: Does the integration task affect EFL learners' products of text comprehension?

Regarding RQ6-1, Experiment 4 suggested that EFL learners had difficulty integrating information across paragraphs although they had been given instructions on the superordinate inference task. The integration task in the present study were directly aimed at such processes, which would lead to success in superordinate inference generation.

Regarding RQ6-2, it can be predicted that the reading task would alter the learners' allocation of cognitive resources during reading as found in previous research (Horiba, 2013; Magliano et al., 1999; van den Broek et al., 2001). For example, the integration task used in this experiment would induce EFL learners to engage in more integrative processes (e.g., inference generation) and more active responses to the text (e.g., reaction, evaluation) compared to when they were told to read the text only for comprehension. On the other hand, EFL learners had difficulty controlling the cognitive resource allocation according to the reading goals given by the instructions (Horiba, 2000). In that case, the learners' think-aloud comments would not be influenced by the task instructions.

Finally, regarding RQ6-3, the task instructions in Experiment 5 did not affect text comprehension after reading. It was suggested that when the task instructions did not require learners to engage in specific types of processing during reading, the task would not affect text comprehension. Previous studies also demonstrated that when learners were given task instructions that required learners to alter specific types of processing during reading, the task instructions had a positive influence on text comprehension (Horiba, 2000; Nahatame, 2014). As the integration task used in the present study required the participants to reflect on earlier information (i.e., information in the previous paragraphs) that is no longer kept in their working memory, it seems that the task instructions will be more concrete and specific for learners. It

can be predicted that the integration task would facilitate the learner's text comprehension.

4.3.2 Method

4.3.2.1 Participants

The participants were 23 Japanese undergraduate and graduate students (7 female, 16 male) who were majoring in a variety of fields, such as social studies, engineering, biology, and medical science. All the participants had studied English for more than six years in Japan, and their English proficiency levels ranged from intermediate to advanced according to their self-reports. A small sample size was chosen so that each individual's data could be examined in detail.

4.3.2.2 Materials

Two expository texts used in Experiments 4 and 5 were also used as the experimental passages in this study. This study only adopted the Implicit version of each text as in Experiment 5.

4.3.2.3 Procedure

Task instructions

To investigate the effects of integration task, the present study set the factor of reading task as a within-participant factor. Each participant read two texts under the two reading conditions: the first text in the Control condition, and the second text in the Task condition. If the participants read the first text in the Task condition and then read the second text in the Control condition, the task instructions given first would affect the second reading. Therefore, so that the task instructions given in the first reading would not affect the processes and products of the second reading, the order of the reading conditions was fixed. The assignment of the two

texts (i.e., the GGW text and NS text) to the reading conditions was randomly counterbalanced because Experiment 5 confirmed that the recall rates of the two texts did not differ.

In the Control condition, the participants were only told to read the text to answer the comprehension questions after reading as in Experiment 5. This instruction would not alter any strategic processing for generating superordinate inferences because they were not given any information about the superordinate inference task.

In the Task condition, participants were asked to read the text to comprehend the author's message and state this overall message aloud after reading each paragraph (i.e., the superordinate inference task). The participants were required to read under the following instruction: "Please read the text in order to understand the author's message conveyed through the overall text, rather than one portion of the text. You are going to state the possible message based on multiple paragraphs, once after reading each paragraph. Therefore, you will state the message a total of four times."

The instructions in the Task condition required them to not only summarize each paragraph, but to integrate information across paragraphs to construct globally coherent representations of the text. Thus, when the participants finished reading the first paragraph, the author told them to answer the writer's message that was conveyed through the first paragraph; then, when they finished reading the second paragraph, the author told them to answer the writer's message by integrating information across the first and second paragraphs. The author gave same instructions to the participants in the third and fourth paragraphs; therefore, the participants repeatedly retrieve previous information, which is stored in their long-term memory, to integrate information distributed across paragraphs. The participants in both conditions were not informed that they would complete the recall task after reading the text.

Experimental sessions

The present study included four sections: (a) practice session, which aimed to accustom the participants to the think-aloud procedure; (b) reading the two experimental passages with a think-aloud task in the Control and the Task conditions; (c) an interference task, which intended to avoid the recency effect on the recall task; and (d) a written recall task. Figure 4.4 illustrates the order of each session.

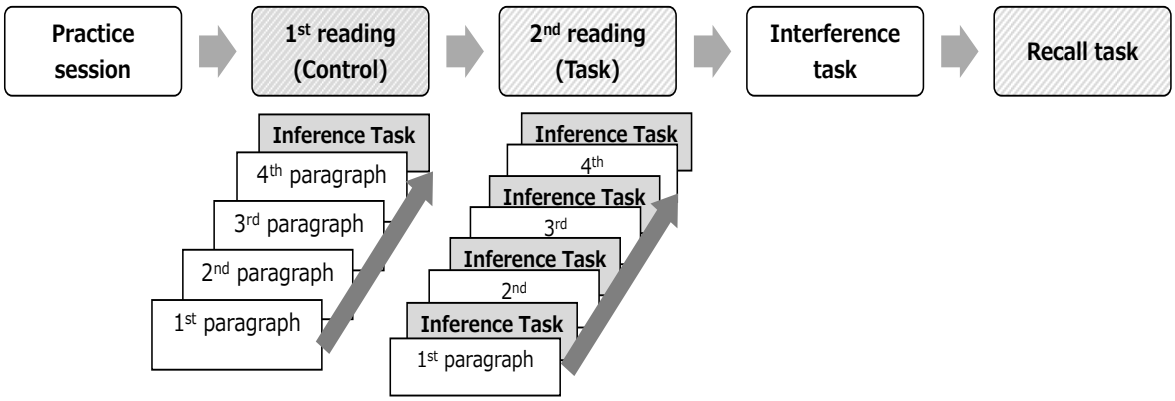

Figure 4.4. The experimental session procedure in Experiment 6.

First, a detailed explanation of the think-aloud task was given in the participants' L1. The think-aloud task required the participants to verbalize in their L1 whatever thoughts came into their minds as they read each sentence. The passages were presented on a computer screen one paragraph at a time. There was a black circle "●" at the end of each sentence to remind the participants to verbalize their thoughts. After reading the paragraph, the participants were asked to press the Enter key to go on the next paragraph. All verbal reports were recorded using an IC recorder. Before reading the two experimental passages, the participants read a practice passage similar in length to the experimental passages, without being given any specific task instructions.

After the practice session, they thought aloud about the first text under the Control condition. In this condition, although the participants were not informed that they would have

the superordinate inference task after reading the text, they completed a superordinate inference task in which they stated aloud the message conveyed through the overall expository texts in one sentence in Japanese after reading the text. After the first reading session, they then thought aloud about the second text under the Task condition. Before this section, they were given detailed instructions about this condition.

Between the think-aloud task and post-reading recall task, participants were asked to read five other English passages as filler passages and answer comprehension questions about them (25 min). This section was intended to avoid the recency effect in the recall task. The passages and questions were the same as the proficiency test used in Experiments 2, 3, and 5. After this interference task, the participants began a written recall task in which they wrote down, in Japanese, all of what they remembered. A time limit was not set so that the participants could recall as much information as possible.

4.3.2.4 Scoring and Analysis

Superordinate inference task

To investigate the relationships between the reading condition and superordinate inference generation, the answers for the superordinate inference task were compared between the Control and the Task conditions. Although the participants answered the writer's message four times in the Task condition, only the final answer (i.e., the answer after reading the fourth paragraph) was analyzed in order to compare the answer with that of the Control condition.

For the superordinate inference task, readers' answers were evaluated based on the same criteria as Experiment 5. The learners who could appropriately describe the superordinate propositions were categorized as "successful," while those who could not were categorized as "failed." Based on the criteria, 30% of the data were randomly selected and scored by two raters,

with an agreement rate of 80.00%. Disagreements were resolved through discussion, and the researcher scored the remaining data alone.

Think-aloud task

To score the participants' think-aloud protocols, Horiba's (2013) framework was adapted because it is intended for Japanese EFL learners and the theoretical interest was similar to the present study (i.e., the effects of task instructions on learners' reading processes). Based on the participant's protocols, some categories were combined and others were deleted. First, verbal protocols were transcribed and then parsed into clauses. Each clause was categorized into one of the following ten categories: (a) *word analysis*, (b) *sentence analysis*, (c) *paraphrase*, (d) *backward inferences*, (e) *predictive inferences*, (f) *association*, (g) *evaluation*, (h) *reaction*, (i) *self-monitoring*, and (j) *other*.[6] The definition of these categories are shown in Table 4.13, and the examples of protocols are shown in Appendix 7.

[6] *Superordinate inference* was originally set as a subcategory of inference. However, all the participants produced comments on superordinate inferences in superordinate inference task rather than in think-aloud comments during reading. Therefore, the comments about superordinate inference were not included in subsequent analysis of think-aloud protocols.

Table 4.13

Categories of Think-Aloud Protocols in Experiment 6

Process level	Category	Definition
Analysis	Word analysis Sentence analysis	The reader attempts to analyze the formal or semantic features of a word, phrase, and sentence, including L1 translation.
Paraphrase		The reader attempts to paraphrase the expression in the text to enhance his/her understanding.
Inference	Backward	The reader generates an inference that is intended to explain the contents of the current sentence by connecting it to prior text or on the basis of general knowledge.
	Predictive	The reader anticipates something about what will occur in the incoming text.
Reader response	Association	The reader generates an inference that is brought to mind by the text that is not intended to enhance the understanding of the textual information.
	Evaluation	The reader makes a comment or states an opinion about the text that is evaluative.
	Reaction	The reader makes a comment to react, often emotionally, to the text.
Self-monitoring		The reader makes a comment about the degree of his/her own comprehension or use of a reading strategy.
Other		The reader comments on things that are not directly related to his/her comprehension of the text.

Word and sentence analysis were subcategories of *analysis* where the participant

attempted to analyze the form or meaning of each piece of information at the surface memory level. In *paraphrase*, the participant tried to paraphrase the expression in the text in order to enhance his/her own understanding. Backward and predictive inferences were subcategories of *in-text inference* where the reader engaged in relational and integrative processes based on the context in the passage or the reader's background knowledge, which leads to constructing coherent representation of the text. On the other hand, association, evaluation, and reaction were considered as *reader response*. Although reader response itself is not intended to enhance the understanding of the text information, they can be regarded as strategies to actively understand the writer's message through thinking about the relationships between the writer, the text, and the readers themselves. In regard to *self-monitoring*, the participant made a comment about the degree of his/her comprehension or use of a reading strategy.

One-fourth of the participants' protocol data were scored by two raters. The agreement between the two raters was 83.65%. Disagreements were resolved through discussion, and the author scored the remaining data alone. The frequency and percentage of think-alouds in each category were calculated.

Written recall task

On the basis of the division of idea units conducted in Experiment 5, 30% of the recall data were randomly selected and scored by two raters, separately. If two-thirds of the information in the IU was reproduced in a participant's recall protocol, one point was given to that IU. The agreement between the two raters was 95.07%. Disagreements were resolved through discussion, and the author scored the remaining data alone. The recall production rate was calculated and compared by reading conditions.

4.3.3 Results

4.3.3.1 Superordinate Inference Task

The participants' answers were categorized into *successful* and *failed* in order to examine how many learners in each reading condition succeeded with superordinate inference generation. Cross-tabulations for each passage were used to examine the relationship between success (i.e., successful and failed) and reading conditions using the data shown in Table 4.14. Two-way chi-square tests showed that the number of learners who succeeded with inference generation significantly differed by reading condition for both texts. Although the effect size was small to medium, the participants were likely to succeed in superordinate inference better when they read the expository text in the Task condition compared to when they read the text in the Control condition.

Table 4.14

Number of Successful and Failed Participants in the Superordinate Inference Task by Reading Condition in Experiment 6

	Control		Task				
	Successful	Failed	Successful	Failed	χ^2	*p*	φ
GGW text	9	14	15	8	2.18	.008	.261
NS text	12	11	18	5	3.45	.006	.274

4.3.3.2 Think-Aloud Task

The number of think-aloud comments for each category by reading condition is presented in Table 4.15. Given that the number of sentences included in each text was 21 for the GGW text and 24 for the NS text, it should be noted that the total number of comments produced by the learners were not so large.

In order to investigate how the learners' cognitive resources were allocated to each process level during reading, the proportion of think-aloud comments were calculated by five process levels. Table 4.16 summarizes the results. As the data in the table show, the participants in both conditions frequently produced lower-level processing, such as word and sentence analyses (about 80% in both conditions), and fewer inferences and responses.

Table 4.15

Number of Think-Aloud Comments for Process Level and Category by Task Condition in Experiment 6

		Control		Task	
Process	Category	*M*	*SD*	*M*	*SD*
Analysis	Word	0.87	1.10	1.96	1.66
	Sentence	28.91	10.91	28.30	9.98
Paraphrase		1.74	2.12	2.22	2.49
Inference	Backward	0.78	1.31	1.13	1.74
	Predictive	0.26	0.62	0.35	0.71
Response	Association	0.00	0.00	0.26	0.62
	Evaluation	0.17	0.49	0.43	0.66
	Reaction	0.17	0.58	0.09	0.29
Monitoring		3.04	3.04	2.52	3.31
Other		0.04	0.21	0.17	0.65
Total		36.00	13.44	37.43	13.95

In order to examine the effects of task instructions on readers' allocation of cognitive resources, a *t*-test was conducted on the proportion of five process level (i.e., analysis,

paraphrase, inference, response, and monitoring). The sample size was so small that the analyses were conducted on each process level respectively. Figure 4.5 shows the production rate of think-aloud comments for process level. Generally, the results showed that the proportion of think-aloud comments did not largely differ between the Control and Task conditions. Specifically, although a significant difference was found in monitoring, the effect size was small, $t(22) = 2.33$, $p = .029$, $d = 0.33$. Regarding other categories, there were no significant differences between reading conditions (ps > .050), and the effect sizes were also small. These results suggested that learners did not change their resource allocation according to task instructions.

Table 4.16

Proportion of Think-Aloud Comments for Process Level by Reading Condition in Experiment 6

	Control		Task	
Process	*M*	*SD*	*M*	*SD*
Analysis	83.97	11.66	83.13	13.98
Paraphrase	4.29	4.92	5.11	5.94
Inference	2.17	3.47	3.05	4.24
Response	1.02	3.27	2.29	4.78
Monitoring	8.39	7.86	5.99	6.79

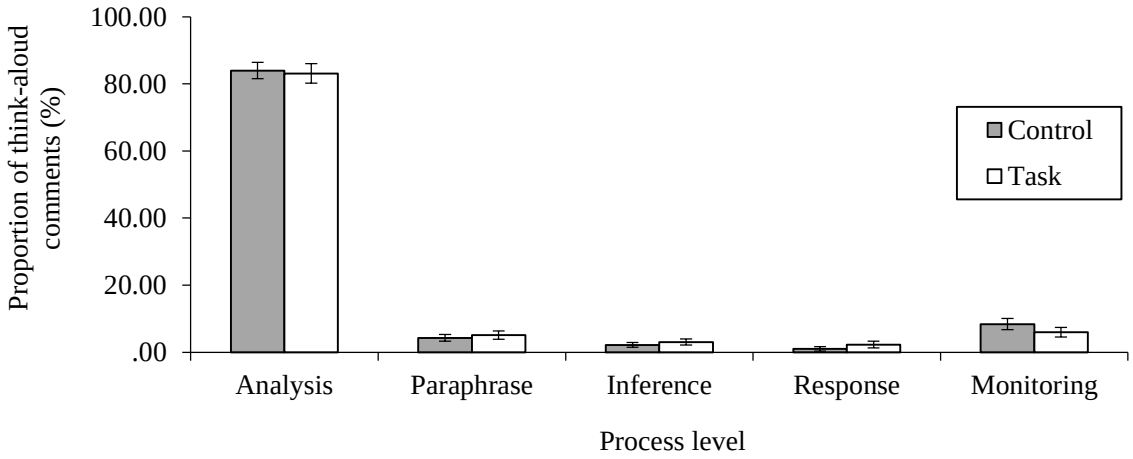

Figure 4.5. Proportion of think-aloud comments for each process level in Experiment 6. Error bars represent standard errors.

4.3.3.3 Written Recall Task

Table 4.17 shows the recall production rate by reading conditions. The results of a *t*-test showed that the difference between the reading conditions was statistically significant, and the effect size was large, $t(22) = 5.24$, $p < .001$, $d = 0.93$. It was indicated that the participants recalled more information in the Task condition than in the Control condition (see also Figure 4.6), suggesting that the reading task aimed at integration of information across paragraphs effectively promoted readers' text comprehension.

Table 4.17

Mean Production Rates With Arcsine Transformation in the Recall Task in Experiment 6

	M	95% CI	*SD*
Control	29.94	[27.29, 32.59]	6.49
Task	35.94	[33.26, 38.62]	6.37

Note. CI = confidence interval.

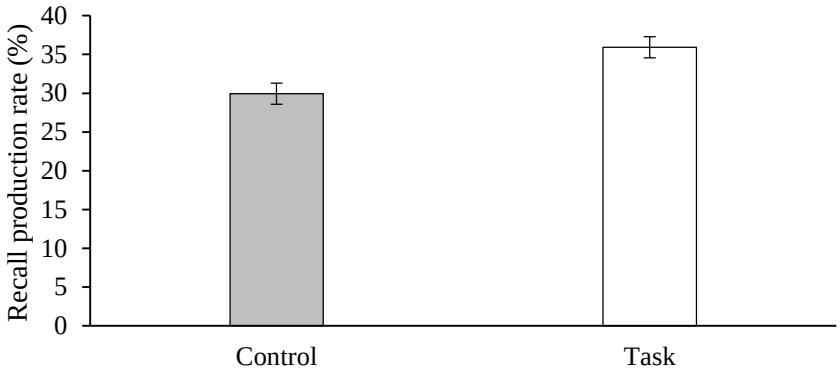

Figure 4.6. Total recall production rates (%) by reading condition in Experiment 6. Error bars represent standard errors.

4.3.3.4 Relationship Between Processes and Products

To examine the relationships between the processes and products of reading comprehension, the correlations between think-aloud comments and recall production rate were tested. Table 4.18 shows the correlation between think-aloud comments and recall production rates.

The results indicated that the correlation patterns were different between reading conditions. Specifically, in the Control condition, word analysis negatively correlated with recall production (r = –.478). This result indicated that the participants who devoted more cognitive resources to word analysis were likely to recall less information in the text. Moreover, the correlation between all inference categories (i.e., backward, predictive, and total) had moderate positive correlations with recall production rate in the Control condition. In the Task condition, on the other hand, only the predictive inference and recall production rate were positively correlated, indicating that the participants who generated more predictive inferences recalled the text better.

Table 4.18

Correlations Between Each Category in Think-Aloud in Experiment 6

		Condition			
		Recall in the control condition		Recall in the task condition	
Process level	Category	*r*	*p*	*r*	*p*
Analysis	Word	–.478	.021	–.202	.356
	Sentence	.042	.848	.020	.929
	Total	–.118	.591	–.069	.756
Paraphrase		.027	.902	–.075	.735
Inference	Backward	.527	.010	.265	.222
	Predictive	.465	.025	.441	.035
	Total	.561	.005	.371	.082
Response	Association	NA	NA	.182	.405
	Evaluation	–.239	.273	.051	.817
	Reaction	–.288	.183	–.074	.738
	Total	–.270	.212	.066	.765
Monitoring		.054	.808	.008	.970
Other		–.315	.143	–.306	.156

Note. NA = not applicable. Comments on association in the control condition were not found.

These results provide evidence that the reading task had a different influence on the processes and products of comprehension. In the Control condition, some processes during reading, such as word analysis and inference generation, did directly relate to products measured by a recall task. In the Task condition, on the other hand, only predictive inference generation related to the products. These differences between conditions are further discussed

later in section 4.3.4.

4.3.4 Discussion

Effects of integration task on superordinate inference generation (RQ6-1)

As well as Experiment 5, some previous studies revealed that L2/EFL readers had difficulty integrating globally distributed information (e.g., Morishima, 2013; Ushiro et al., 2008). Additionally, in the L1 reading research, inference generation in expository reading can be difficult (e.g., Noordman et al., 1992; Singer et al., 1997; Wiley & Myers, 2003) due to the lack of prior knowledge and its complexity as reviewed in section 2.1.2. Therefore, it was predicted that the task instructions in which learners were required to integrate information distributed across paragraphs would facilitate learners' generation of superordinate inferences in expository reading. The results of the superordinate inference task revealed that the task instructions given in the present study effectively facilitated superordinate inference generation.

According to the models of reading comprehension, such as the landscape model (van den Broek et al., 1996), the extent to which readers engage in processes to construct a coherent representation depends on the reading goal given by the task instructions. To construct a globally coherent representation, readers must link currently processing elements and much earlier elements that are no longer kept in their working memory and/or to relevant world knowledge. However, while information, which is locally located with currently processing information can be activated easily and spontaneously, the information from a much earlier text can be strategically retrieved (e.g., McKoon & Ratcliff, 1992).

The reason why the integration task in Experiment 6 facilitated superordinate inference generation whereas the task instructions in Experiment 5 did not might be attributed to how strongly the task induced the learners to reinstate and retrieve the earlier text elements stored in their long-term memory. In Experiment 5, the participants were just told to think about the

writer's message conveyed through the overall text, but were not told to engage in the retrieval processes during reading. The integration task in Experiment 6 required readers to repeatedly access previous information, which is stored in their long-term memory. Therefore, the learners constructed a globally coherent representation of the expository text more easily in the Task condition compared to in the Control condition.

Although the analysis of superordinate inference task in the Task condition was mainly conducted on the final answer (i.e., the answer after reading the fourth paragraph) in section 4.3.3.1, some comments in the superordinate inference task in the Task condition can be regarded as retrieval and integration processes induced by the task. For example, one participant stated the author's message as "The desert's growth is a serious problem in Africa." after reading the first paragraph of the GGW text, and then answered "To solve the problem of the desert's growth, African countries started planting trees on a large scale." "Although the project of planting trees aims to stop the desert's growth, some people concern the financial and biological problems." after reading the second and the third paragraphs, respectively. Finally, the participant generated appropriate superordinate inference "The Great Green Wall is a good solution for the desert's growth in Africa." after reading the fourth paragraph. Underlined parts can be regarded as information retrieved from the previous paragraphs which is already stored in his long-term memory. As the participants in the Task condition were told to answer the author's message based on not only information in the paragraph they had just read but also information across paragraphs, most of the participants tried to answer the author's message based on the multiple paragraphs. The results suggested that Japanese EFL learners attempted to retrieve previous information and maintain the coherence between previous and current information while engaging the integration task, which led to superordinate inference generation.

Thus, it can be concluded that superordinate inference generation is important for

constructing globally coherent representation of expository texts but that it is strategic rather than spontaneous processes for L2 text comprehension that are influenced by the reading task aimed at integrating information across paragraphs.

Effects of integration task on EFL learners' processes of reading comprehension (RQ6-2)

The effects of integration task on learners' reading processes were tested by the think-aloud task. Two possibilities about how the reading task would affect the processes during reading were considered. First, the integration task could alter the learners' allocation of cognitive resources during reading. For example, the task used in this experiment would induce EFL learners to engage in more integrative and active processes compared to when they were told to read the text just for comprehension. In this case, it would be expected that learners would produce more comments on inference generation and responses to the text in the Task condition compared to in the Control condition. Second, EFL learners had difficulty controlling the cognitive resource allocation according to the reading goals given by the integration task. In this case, the pattern of think-aloud comments would not be different between the two conditions.

The results showed that the proportion of think-aloud comments did not largely differ between the Control and Task conditions. Although a significant difference between the two conditions was found in the process of monitoring, the effect size was small. The reason the learners' processes during reading did not differ according to task instructions can be explained by the following two points. First, they devoted too many resources to lower-level processing, such as word and sentence analyses, regardless of the reading condition. As many as about 80% of the comments produced by the participants were lower-level processing, such as word and sentence analyses. Many previous studies, which adopted the think-aloud task also suggested that L2 readers were likely to use many cognitive resources to lower-level linguistic processing

compared to L1 readers (Horiba, 1996, 2000). Shimizu (2015) also suggested that L2 readers (especially lower-proficiency readers) tended to allocate more cognitive resources to lower-level processing in expository reading compared to in narrative reading. The present study was consistent with these studies. Because only a small amount of cognitive resources was available, the learners could not allocate their resources to engage in relational and integrative processing in accordance with the task instructions. The results suggested that nonnative readers were likely to primarily engage in the language competence-based processing as in Horiba (2000), rather than the task-based processing.

In addition to the think-aloud comments, the time spent on reading experimental passages with think-alouds was recorded using SuperLab 4.5 software. According to Magliano et al. (1999), if reading time for the Task condition is longer than the Control condition, it means that the reading task requires additional cognitive effort over the Control condition. The total reading time (sec) for the Control condition was M = 504.63, SD = 233.51 and the reading time for the Task condition was M = 535.25, SD = 218.58. The reading time for the Task condition was likely to be longer than that for the Control condition; however, a paired t test showed no significant differences between two conditions due to the small number of participants, t (22) = –1.21, p = .241, d = 0.13. Although the time spent on reading the passage with think-alouds was not influenced by the reading task, the results of the superordinate inference task suggested that participants engaged in retrieval and integrative processes after reading each paragraph in the Task condition. In other words, while they primarily engaged in language competence-based processing (e.g., word and sentence analyses) during reading the text, they engaged in task-induced additional processing only when they were explicitly told to do so in the superordinate inference task.

Although the processes themselves did not differ between the reading conditions, the standards of coherence (van den Broek et al., 2001), which influence the extent to which readers

engage in processes to construct coherent representation, might have been affected by the task instructions. Taking inference generation as an example, while the learners were asked to just read the text to understand the content in the Control condition, some of them were likely to engage in relational and integrative processes in order to enhance their comprehension of the text. In the Task condition, inference generation during reading might aim to construct the writer's message conveyed through the overall text rather than to comprehend the text itself. Therefore, this possibility will be further discussed in the later section, based on the correlation between the think-aloud data and recall production rate.

Effects of integration task on EFL learners' products of reading comprehension (RQ6-3)

In order to investigate the effects of integration task on the products of reading comprehension, the learners' recall production rate was analyzed. While many L2 studies examined the effects of task instructions on text comprehension, only a few studies found positive effects on text comprehension (e.g., Horiba, 2000; Nahatame, 2014). As reviewed in section 2.3.2, these studies suggested that when learners were given task instructions that required them to alter specific types of processing, it would lead to a positive influence on text comprehension. As the task instructions in the present study (i.e., to integrate information distributed across paragraphs) were more concrete and specific for learners compared to that of Experiment 5 (i.e., to think about the writer's message conveyed through the overall text), it was predicted that the task instructions would facilitate the learner's text comprehension. In accordance with the prediction, the results of the recall task demonstrated that the participants recalled more information in the expository text in the Task condition compared to the Control condition. The results suggested that the reading task aimed at the integration of information across paragraphs effectively promoted readers' text comprehension.

As along with the specificity of task instructions, there are other possible reasons for this

positive effect of reading tasks on text comprehension. First, the participants' engagement in retrieval and integrative processes in the superordinate inference task after reading each paragraph contributed to constructing robust mental representations. Previous studies demonstrated that cognitively demanding tasks may have had negative effects on reading comprehension (e.g., Horiba, 2002); in this study, however, text processing and task-induced additional processing were likely to occur alternately. Specifically, in the Task condition, while learners engaged in lower-level linguistic processing while reading the text, they engaged in higher-level conceptual processing as they completer the superordinate inference task. As a result, the integration task facilitated both learners' text comprehension and superordinate inference generation.

The second reason is that the superordinate inference task helped readers to organize subordinate information in the expository text. Previous studies demonstrated that when the superordinate macroproposition is explicitly stated as a title or topic sentence of the text, readers used it as an organizational cue in order to understand the gist of a text (e.g., Lorch et al., 1993; Sanchez et al., 2001). In the integration task of the present study, the participants inferred a superordinate proposition after reading each paragraph and then may have used it as an organizational cue to process the textual information in the subsequent paragraphs. In other words, generation of superordinate inferences effectively facilitated processing of consequent textual information.

Although the interference task was given to the participants between the think-aloud task and post-reading recall task (see Figure 4.4), a possible argument against this result might be that the higher recall production rate in the Task condition was due to a recency effect because the participants read the passage in the Control condition first, and then the second passage in the Task condition. To test this possibility, 12 Japanese EFL learners, who did not take part in Experiment 6, read both passages under the Control condition. After the same interference task

as in the experiment, they completed the written recall tasks. The results demonstrated no difference between the first (M = 31.14, SD = 4.07, 95% CI [28.84, 33.44]) and the second reading (M = 31.32, SD = 4.73, 95% CI [28.64, 33.99]) in their recall production rate, $t(11)$ = –.121, p = .906, d = 0.04. Thus, the improvement of the recall production rate in the Task condition in the experiment can be attributed to the effects of task instructions rather than the recency effect caused by the presentation order of the texts.

In summary, the task instructions aimed at integration of information across paragraphs had positive effects on the products of reading comprehension. As in Horiba (2000), which examined the effects of task instructions for facilitating connecting the current sentence and prior/later text (i.e., relatively local range of the text), the present study also demonstrated the facilitative effects of task instructions, focusing on integrating information distributed more globally in the text. Giving task instructions that have participants focus on constructing globally coherent representations of the text enabled learners to construct more stable and elaborative text representations in their memories, which led to better comprehension of the text.

Combining the results of the think-aloud protocols and the recall task, the present study demonstrated that while the learners' processes during reading did not change according to the reading task, the products of comprehension were influenced by the task. This suggests the possibility that although the overall pattern of think-aloud data did not differ between reading conditions, there might be some qualitative differences between when the learners were given the integration task and when they were not. In other words, the learners' standards of coherence might have been affected by the reading task (van den Broek et al., 2001). Therefore, the following section discusses the relationships between processes and products, focusing on the correlations between think-aloud comments and recall production rates.

Relationship between processes and products of reading comprehension

The relationships between the processes and products of reading comprehension were examined by the correlation analysis between think-aloud comments and recall production rate. As stated above, although the effects of the integration task were not found in the proportion of think-aloud comments, there might be some qualitative differences between when the learners were given task instructions and when they were not.

Some L2 studies directly examined the relationships between the processes and products of text comprehension (e.g., Horiba, 2000, 2013). In Horiba (2000), although the relationships between the processes and the products were not significantly correlated, the descriptive statistics showed that in the read-freely condition, which is the similar condition to the Control condition of the present study, L2 readers who produced more comments on backward and forward inferences in their think-alouds tended to perform better in the recall task compared to those who did not. Meanwhile, when the task instruction (i.e., read for coherence) was given to L2 readers, such a relationship was not found. She suggested that when the L2 readers were given the instruction to pay attention to the relationships between the current sentence and prior/later text, they made extra efforts for such processing. Consequently, the task instructions might weaken the relationships between the processes and products of L2 text comprehension. In another study, Horiba (2013) found no significant correlations between the processes and products of comprehension, suggesting that L2 text comprehension was "complex and not straightforward."

The present study results found a similar pattern with Horiba (2000). The results showed that the reading task partly affected the relationships between the processes and the products of L2 text comprehension. In the Control condition, some processes during reading, such as word analysis and inference generation, did directly relate to text comprehension measured by a recall task. In the Task condition, only predictive inference generation related to text comprehension.

These findings can be accounted for by the following two possibilities. The first, based on Horiba (2000), is that the participants in the Task condition could make extra efforts toward relational and integrative processes. However, considering the proportion of think-aloud comments shown in Table 4.15, the comments on such processing were very few. Therefore, this possibility may be rejected in the present study. The second possibility is that although the proportion of the think-aloud comments were similar in the two conditions, the reason why the readers engaged in each process might differ according to the reading goals given by the task instructions. Specifically, in the Control condition, the participants were told that they would answer comprehension questions after reading the text. Such instructions might induce learners to engage in relational and integrative processing in order to improve their comprehension of text content, consequently leading to significant correlations between processes and products. In the Task condition, they were told to integrate information across paragraphs in order to answer the superordinate inference task rather than to comprehend the text itself. Therefore, the processes they engage in the Task condition might aim at superordinate inference generation.

Although the present study did not directly investigate the effects of task instructions on the learners' reading goals as in Experiment 3, the detailed observation of think-aloud protocols may provide some clues to confirm the possibility. For example, the following is an example of a participant's processes when reading the third paragraph of the NS text in the Control condition.

> Mosquitoes laid eggs in the water … and wrigglers hatched from the eggs … The wrigglers have lived in the water for one or two weeks. First of all, it has no damage to the environment…. It seems logical because this method did not use chemicals. (Author's translation)

In this protocol, the underlined parts were categorized as inferences because they were

not explicitly stated where the participant was currently reading. Specifically, the word "wrigglers" was derived from the participant's own prior knowledge that "wrigglers" hatched from mosquito eggs. In addition, although the information about chemicals was not stated explicitly in the third paragraph, "because this method did not use chemicals" was derived from the first paragraph, which the participant had already read (i.e., *they* [the chemicals] *are often bad for the environment*). This example implied that the participant generated inferences that contributed to connecting information beyond the paragraph (i.e., global inference), even in the Control condition without specific task instructions. In fact, the recall production rate of this participant was more than average in the Control condition.

Regarding the think-aloud protocols in the Task condition, let us consider the following data when reading the second paragraph of the GGW text.

> The trees will be a home for plants and animals … and the vegetation will recover…. Then, because the trees also reduce CO_2, ... it will become a measure against global warming. Also, it contributes to the reduction of greenhouse gas; and, accordingly, the desertification will slow down. (Author's translation)

The underlined parts were not stated explicitly in the text, so they were categorized as backward or predictive inferences; for example, "the recovery of vegetation," "a measure against global warming," and "the reduction of greenhouse gas" were not stated explicitly in the text, and the participant predicted these as consequences of the event based on the text information. In the Task condition, participants were induced to read the text in order to complete the superordinate task after reading each paragraph. This means that the readers' main goal for reading was to answer this task rather than understand the text itself. Indeed, after reporting this protocol, the participant answered, "The Great Green Wall was an effective

measure to stop the desert's growth, and it has many advantages" in the superordinate inference task. This answer directly reflected the inferences generated by the participant during reading. After reading only the first and second paragraphs of the text, the participant generated predictive inferences to construct the author's message conveyed through the overall text.

The examples above supported the notion that the processes they engaged in for the Task condition aimed at superordinate inference generation but not at text comprehension itself, which lead to insignificant correlations between the processes and products of text comprehension in the Task condition. However, despite such indirect relationships between processes and products, the learners' recall performance was better in the Task condition compared to in the Control condition.

The reason for this improvement of recall production rates in the Task condition might be attributed to the engagement in the superordinate inference task immediately after reading each paragraph, rather than the processes they engaged in during reading. In the Task condition, the learners answered the author's message conveyed through the overall text after reading each paragraph; therefore, repeated engagement in such a task might facilitate learners' construction of a globally coherent representation of the text, leading to better text comprehension compared to reading in the Control condition.

In summary, the relationships between processes and products differed according to the task instructions, indicating that why the learners engaged in each process depended on the given task. Specifically, when the learners were asked to just comprehend the text, they generated some inferences in order to improve their text comprehension after reading. After being asked to integrate information across paragraphs to comprehend the writer's message, the learners were likely to engage in conceptual processes during reading in order to complete the task itself rather than to comprehend the text. These results indicated that while the task instructions, which aim to have learners construct a globally coherent representation of the text,

did not directly alter the processes of reading, learners constructed coherent text representations through engaging in the superordinate inference task, consequently leading to better products of text comprehension.

4.3.5 Conclusion of Experiment 6

The purpose of Experiment 6 was to investigate the effects of integration task that induced readers to construct global coherence of the text, focusing on the processes, products, and relationships between the processes and products of Japanese EFL learners' text comprehension. Based on the results of Experiment 5, the present study gave the reading task, which aimed to facilitate the integration of information distributed across paragraphs. Specifically, the participants in the Task condition were asked to answer what the writer intended to convey each time they finished reading each paragraph. The findings of this study can be summarized as the following points.

First, the integration task effectively facilitated Japanese EFL learners' superordinate inference generation (RQ6-1). The instructions in the present study required learners to repeatedly access previous information, which is stored in their long-term memory, which helped them to construct a globally coherent representation of the expository text more easily in the Task condition compared to in the Control condition. Combining the results of Experiments 4 and 5 and the present study, it can be concluded that superordinate inference generation for expository texts is a strategic rather than spontaneous processes for EFL text comprehension, which is influenced by the reading task.

Second, the results of the think-aloud data demonstrated that learners did not alter their cognitive processes according to the reading tasks in expository reading (RQ6-2). Regardless of the reading condition, they tended to allocate much of their cognitive resources to lower-level linguistic processing, such as sentence analysis, while the higher-level conceptual

processing, such as inferences and responses to the text, were limited. On the other hand, the results of the written recall task indicated that the learners' text comprehension was facilitated by the integration task (RQ6-3).

The third and most interesting finding was that the relationships between processes and products differed according to the reading task. Specifically, the processes measured by the think-aloud task were significantly related to the products tested by the recall task in the Control condition, whereas such relationships were partly found in the Task condition. The detailed observation of the think-aloud data indicated that while the learners in the Control condition engaged in processes in order to improve their text comprehension, they engaged in their processing in order to answer the superordinate inference task in the Task condition. These results suggested that the integration task did not directly change the learners' processes during reading; rather, it influenced the learners' goals for reading, which resulted in better text comprehension in the Task condition.

In sum, although the previous studies and Experiments 4 and 5 assumed that Japanese EFL learners had difficulty constructing global coherence of expository texts, the integration task helped learners to build globally coherent representations of the expository texts. These results supported the notion that superordinate inference in expository reading was not spontaneous processing but rather strategic processing, which is influenced by reading goals given by the task. Moreover, such strategic processing enabled learners to construct more stable and elaborative text representations in their memories, which led to better performance in the post-reading task, such as the written recall task.

Chapter 5

General Discussion

In order to examine the construction of globally coherent mental representations among Japanese EFL learners, the current research conducted a total of six experiments. First, Study 1 of this dissertation conducted three experiments (Experiments 1 to 3) to examine thematic inference generation in narrative reading. Next, Study 2 conducted three experiments (Experiments 4 to 6) to investigate superordinate inference generation in expository reading. The following sections summarize the main findings of this research and generally discuss the six experiments from multiple perspectives.

5.1 Building Globally Coherent Mental Representations of Narrative Texts Through Thematic Inference Generation

Study 1 involved conducting three experiments (Experiments 1 to 3) in order to answer the following six research questions regarding Japanese EFL learners' thematic inference generation in narrative reading:

RQ1-1: Do Japanese EFL learners understand implicit themes in narrative texts by generating thematic inference?

RQ1-2: Does the generation of thematic inference differ from other inference types?

RQ2-1: Do Japanese EFL learners strategically generate thematic inferences when instructed to think about the overall message conveyed by the writer?

RQ2-2: Do task instructions aimed at thematic inferences affect Japanese EFL learners' text comprehension?

RQ3-1: Do task instructions aimed at thematic inference generation change EFL learners' reading goals measured by a questionnaire?

RQ3-2: Do task instructions aimed at thematic inference generation change EFL learners' cognitive processes measured by a think-aloud task?

First, Experiment 1 examined whether Japanese EFL learners understand implicit themes in narrative texts by generating thematic inference (RQ1-1) and whether the generation of thematic inference differs from other types of inferences (RQ1-2). This experiment manipulated the explicitness of thematic statements in narrative passages and adopted an inference verification task where participants evaluated if verification statements can be inferred from the passages using both yes/no responses and those on a 5-point Likert scale. The results of the inference verification task indicated that the participants evaluated explicit themes as more valid than implicit themes, suggesting that it can be difficult for Japanese EFL learners to understand implicit themes through inference generation. In addition, it was demonstrated that thematic inference was generated more than emotional inference, while thematic inference was similar to goal, action, and state inferences.

Experiment 1 did not require readers to engage in a specific reading strategy; therefore, their reading goal might be to construct local coherence of the mental representation rather than building global coherence. Consequently, in order for learners to construct global coherence of narrative texts, educational interventions are required. In addition, thematic inference generation needs to be replicated with a methodology other than the inference verification task because it only required learners to judge the presented statements and did not measure learners' generation of thematic inference content. To address these issues, Experiment 2 investigated the effects of task instructions on thematic inference generation using a thematic inference task wherein participants were required to identify the themes of narrative texts.

In Experiment 2, the following two points were investigated: whether task instructions facilitate thematic inference generation among Japanese EFL learners (RQ2-1), and whether task-induced strategic processing affect text comprehension (RQ2-2). Accordingly, Experiment 2 compared the difference between the Task condition (i.e., reading passages in order to understand the theme conveyed by the writer) and the Control condition (i.e., reading passages for comprehension).

The results of the thematic inference task demonstrated that the Task condition participants performed better than those in the Control condition. In addition, when learners were not given any strategic instructions, they were likely to construct themes that were too narrow and based on local and partial text information rather than overall text. These results suggested that task instructions effectively facilitated learners' thematic inference generation. Additionally, in the written recall task, it was revealed that the participants in the Task condition recalled more information than those in the Control condition. Moreover, detailed analysis of the recall protocols demonstrated that the participants in the Task condition recalled outcome-related information significantly better than those in the Control condition. These results suggested that strategic instructions for thematic inference generation helped learners construct a globally meaningful representation that led to better text comprehension.

Although positive effects of task instructions were found in both thematic inference generation and text comprehension, the reason why such facilitation occurred remained unclear in Experiment 2. Consequently, to clarify the effects of task instructions, it was necessary to examine the on-line processes that EFL readers engage in during reading.

Experiment 3 addressed the following two issues: whether the task instructions aimed at thematic inference generation change EFL learners' reading goals (RQ3-1) and alter EFL learners' cognitive processes (RQ3-2). Experiment 3 provided the same task instructions used in Experiment 2, and measured learners' cognitive processes during reading utilizing a think-

aloud method. In addition, think-aloud data was combined with responses to a brief questionnaire in order to measure the reading goals that learners tried to achieve during reading comprehension.

The questionnaire results revealed that when task instructions were given, participants' reading goals shifted from understanding story characters to building global coherence of texts, understanding causal relationships between textual information, and understanding characters' intentions. On the other hand, think-aloud data demonstrated that the effects of task instructions only appeared in parts of the reading process. Specifically, task instructions reduced the allocation of cognitive resources to lower-level linguistic processing such as analysis of each word and sentence.

Additional analysis of the relationships between L2 reading proficiency and cognitive resource allocation indicated that while less proficient learners were likely to allocate more cognitive resources to lower-level linguistic processing (e.g., analyzing words and sentences) regardless of specific task instructions, instructions enabled proficient learners to allocate more cognitive resources to higher-level conceptual processing (e.g., backward and predictive inferences). Therefore, combining the results of questionnaire and think-aloud tasks leads to the conclusion that, although both proficient and less proficient learners tried to employ different cognitive strategies based on task instructions, only highly proficient learners could partially change their processing to meet reading goals.

Based on this summary of results obtained from Experiments 1 to 3, the following three points should be discussed, respectively: (a) whether thematic inference generation is the result of automatic or strategic processing, (b) the effects of thematic inference generation on text comprehension, and (c) the relationship between cognitive resource allocation and L2 reading proficiency.

Is thematic inference generation the result of automatic or strategic processing in narrative comprehension?

Text comprehension theories have attempted to explain what inferences are made and under what conditions (e.g., Graesser et al., 1994; McKoon & Ratcliff, 1992; Singer et al., 1994). In particular, there is disagreement between the minimalist hypothesis (McKoon & Ratcliff, 1992) and constructionist theory (e.g., Graesser et al., 1994), especially in regard to the generation of inferences that establish global coherence. Specifically, while the minimalist hypothesis claims that thematic inferences are generated only strategically and not automatically, the constructionist theory claims that thematic inferences are spontaneously generated.

In Experiment 1, where the participants were only instructed to respond in the inference verification task, they judged implicit themes as less appropriate than explicit themes. This might be attributed to a tendency for Japanese EFL learners to execute bottom-up lower-level linguistic processes such as encoding each element of textual information to construct local text coherence rather than attending to top-down processing such as comprehending the writer's messages conveyed throughout overall passage. In Experiment 2, while learners in the Control condition were more likely to construct themes that were too narrow and based on part of the narrative text, those in the Task condition performed better on the thematic inference task. Finally, although Experiment 3 did not directly examine thematic inference generation, the think-aloud data indicated that few participants commented on thematic inference during reading.

Although the present study did not primarily focus on on-line generation of thematic inference, these results can lead to the conclusion that thematic inferences were not automatically generated during EFL reading, while they were strategically generated when task instructions were directed at specific reading goals. Likewise, previous studies have reported

that, even in L1 reading, readers only generated thematic inferences when provided with specific reading goals (e.g., Kurtz & Schober, 2001; Seifert et al., 1986). Compared to L1 readers, Japanese EFL learners' language proficiency is limited and they tend to allocate more cognitive resources to lower-level linguistic processing than higher-level conceptual processing (Horiba, 1996, 2000; Morishima, 2013). Therefore, EFL learners generated thematic inferences only when they had specific goals for constructing globally coherent mental representations and engaged in strategic processing to achieve their goals.

However, this study involved relatively limited participants and materials; thus, whether thematic inferences are automatically or strategically generated should be further examined. For example, if participants' L2 reading proficiency had been more varied, an interaction between L2 reading proficiency and task instructions might have been found. In addition, as demonstrated by Zhang and Hoosain (2001), the text factor (e.g., title presentation) could have influenced thematic inference generation. The current study's limitations are discussed later in this chapter.

The effects of thematic inference generation on text comprehension

According the constructionist theory, thematic inference contributes to building global coherence of texts (Graesser et al., 1994); however, whether and how thematic inference generation affects text comprehension remains unclear. Using the written recall task, the present study examined whether strategic instructions related to thematic inference enhanced Japanese EFL learners' text comprehension. Previous L2 studies about the effects of task or strategy instructions on text comprehension demonstrated that text comprehension was facilitated when learners were given instructions that required them to alter specific processing types during reading (e.g., task instructions aimed at a specific type of inference) (e.g., Horiba, 2000; Nahatame, 2014). Conversely, when learners' were not told to use a specific strategy or

processing behavior during reading, clear effects of task instruction were not found in text comprehension (e.g., Horiba, 2013; Yoshida, 2012). Furthermore, task instructions could negatively affect L2 text comprehension task requirements by overloading learners (Horiba, 2002).

Experiments 2 and 3 tested the effects of strategic processing of thematic inference generation on L2 text comprehension by adopting the written recall task. Experiment 2 showed that the participants in the Task condition had better recall than those in the Control condition. Consequently, this finding suggested that text comprehension was facilitated by task-induced strategic processing of thematic inference generation. In addition, it was demonstrated that the facilitative effect of task instructions was especially evident in outcome-related information contained in the narrative passages. Experiment 3 showed similar facilitative effects on the total recall production rate; however, the effects of task instructions facilitated recall production of each story category.

Based on the results of the written recall tasks in Experiments 2 and 3, it can be concluded that task instructions aimed at thematic inference generation helped Japanese EFL learners attend to important elements, which contributed to the construction of a globally coherent and robust representations of narrative texts. According to Trabasso and colleagues (Trabasso & van den Broek, 1985; Trabasso et al., 1989), narrative texts describe the causal relationships among various events; therefore, readers need to understand not only individual text elements but also the causal relationships between these events. Thus, in order to generate thematic inferences, readers need to identify consistency between an action (or goal) and its outcome (e.g., Dorfman & Brewer, 1994; Zhang & Hoosain, 2001). The task instructions given in Experiments 2 and 3 enabled EFL learners to engage in this process, which led them to relate text elements to one another.

The relationship between cognitive resource allocation and L2 reading proficiency

Thus far, previous L1 and L2 reading studies have investigated the relationships between cognitive resource allocation and reader-related factors (e.g., Bohn-Gettler & Kendeou, 2014; Bråten & Strømsø, 2009; Horiba, 1996, 2000; Linderholm & van den Broek, 2002). Most of existing studies reported that readers with higher working memory or reading proficiency could better control their cognitive processes during reading in accordance with task instructions than those with lower working memory capacity and reading proficiency. In addition, it was demonstrated that L2 readers had less flexible control of their reading processes than L1 readers.

In Experiment 2, the results of the written recall task reveled main effects of only L2 reading proficiency and task instructions, and the interaction between these factors was non-significant. Additionally, in Experiment 3, main effects of task instructions were only found for recall production rate and questionnaire responses. On the other hand, correlational analysis of L2 reading proficiency and think-aloud comments suggested some interesting relationships between two factors. Specifically, in the Control condition, the lower the level of learner proficiency, the more cognitive resources were allocated to lower-level linguistic analysis. A similar tendency was also found in the Task condition. Furthermore, in the Task condition, the higher the level of learner proficiency, the more cognitive resources were allocated to higher-level conceptual processing (e.g., backward and predictive inference generation).

In summary, the results of the products of comprehension measured by the recall task suggested that task instructions facilitated EFL learners' text comprehension, regardless of L2 reading proficiency. The narrative passages used in Study 1 simplified low frequency words and complex syntactic construction for participants. As a result, the influence of participants' linguistic knowledge on their text comprehension might have been small. A study by Muramoto (2000) that investigated L2 learners' inference generation in simple narrative passages reported similar results. Specifically, the difference between skilled L2 readers and less skilled readers

responses to simple passages was attributed to how they processed the text rather than their amount of linguistic knowledge. In fact, the think-aloud protocols adopted in Experiment 3 demonstrated that less proficient learners allocated too many cognitive resources to analyzing words and sentences in the simple passages. In contrast, proficient learners assigned their resources to higher-level conceptual processing in order to elaborate and embellish the representations of narrative texts.

Figure 5.1 illustrates the construction of global coherence of narrative texts in Japanese EFL learners' reading comprehension, focusing on the effects of task instructions and L2 reading proficiency on readers' standards of coherence, cognitive processes during reading, and products of text comprehension. When readers were given task instructions for constructing implicit themes, readers tried to maintain the coherence of narrative texts in terms of character's goals and causation. Moreover, they attempted to focus on global coherence of narrative comprehension. By setting their standards of coherence, the allocation of cognitive resources to lower-level processing were reduced while the instructions enabled proficient learners to allocate more cognitive resources to higher-level conceptual processing (e.g., backward and predictive inferences). As a result, task instructions facilitated thematic inference generation and learners' construction of robust mental representations which represented the coherence of character's actions and outcomes.

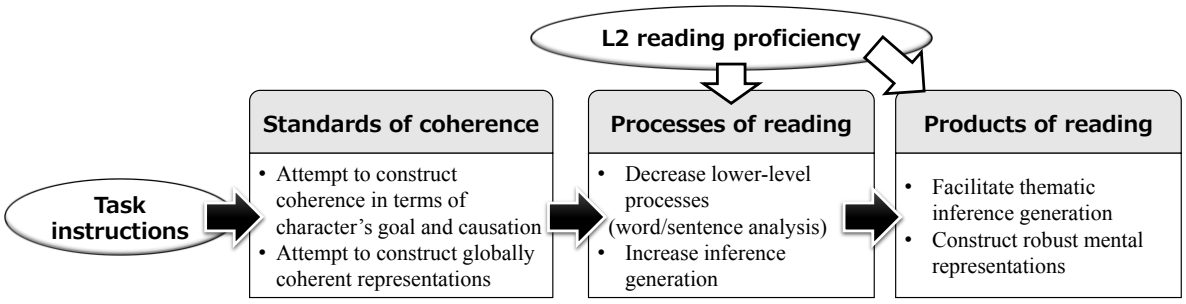

Figure 5.1. Construction of global coherence of narrative texts in Japanese EFL learners' reading comprehension.

Construction of global coherence of narrative texts in Japanese EFL learners' reading comprehension shown in Figure 5.1 can be explained by theories and models regarding reading comprehension and inference generation. The present study concluded that thematic inference was not automatically generated during L2 reading, while it was strategically generated when task instructions were directed at specific reading goals. The present study supported the notion of the minimalist hypothesis, in that thematic inference is only generated strategically and not automatically (McKoon & Ratcliff, 1992).

In addition, the effects of task instructions on inference generation and text comprehension can be accounted for by the landscape model (van den Broek et al., 1996). According to this model, when the activation of concepts by automatic processes are not sufficient for their standards, then readers engage in strategic processes until the levels of comprehension meet those standards. In the case of the present research, when the learners were given task instructions aimed at thematic inference, their reading goals shifted to building global coherence of texts, understanding causal relationships between textual information, and understanding characters' intentions. Although the actual change of the learners' cognitive processes during reading was relatively limited, the higher the level of learner proficiency, the more cognitive resources were allocated to inference generation.

Furthermore, although the present study did not directly investigate the activation of individual concepts during reading, which is assumed in the landscape model, the results of the recall tasks in the present study suggested the possibility that the task instructions altered the activation pattern of individual concepts during reading. Specifically, learners focused on the characters' goals and causal relationships based on their standards of coherence. Consequently, the activation levels of character's actions and outcomes increased because they were strongly related to the characters' goals (Trabasso & Wiley, 2005). As individual concepts were causally related each other, this led construction of coherent and well-organized representations of

narrative texts, which is represented in the improvement of recall production rates.

5.2 Building Globally Coherent Mental Representations of Expository Texts Through Superordinate Inference Generation

Study 2 conducted three experiments (Experiments 4 to 6) in order to answer the following seven research questions regarding Japanese EFL learners' superordinate inference generation in expository reading:

RQ4-1: Do Japanese EFL learners understand implicit superordinate propositions in expository texts by generating superordinate inferences?

RQ4-2: Do mental representations constructed by learners differ according to the explicitness of the superordinate propositions in the text?

RQ5-1: Do task instructions facilitate EFL learners' superordinate inference generation in expository reading?

RQ5-2: Do task instructions aimed at superordinate inference affect Japanese EFL learners' text comprehension?

RQ6-1: Does the integration task facilitate superordinate inference generation?

RQ6-2: Does the integration task affect EFL learners' processes of text comprehension?

RQ6-3: Does the integration task affect EFL learners' products of text comprehension?

Experiment 4 was conducted to answer the following two questions: whether Japanese EFL learners can understand implicit superordinate proposition of expository texts through inference generation (RQ4-1) and whether the presence of superordinate proposition affect learners' mental representations of expository texts (RQ4-2). This experiment manipulated the explicitness of superordinate proposition in expository texts and, as in Experiment 1, adopted

an inference verification task. The results of the inference verification task showed that participants judged implicit superordinate proposition as highly appropriate when explicit, suggesting that superordinate inferences were generated spontaneously rather than strategically without any specific tasks. Furthermore, the explicitness of superordinate proposition affected mental representations. While mental representation of the text is constructed around the superordinate proposition in the Explicit text, the absence of superordinate proposition can prompt learners to instead construct a mental representation based on the local main idea (i.e., the main idea of each paragraph) rather than the hierarchically highest macroproposition.

The results of Experiment 4 suggested that superordinate inference involved spontaneous processing in expository reading. However, it should be noted that the inference verification task required learners to only judge the presented statements. Consequently, they might not activate superordinate inferences until the Consistent statement was presented as the target statement in the verification task. To confirm this possibility, superordinate inference generation was replicated with another methodology in Experiment 5. Moreover, although Experiment 4 suggested the explicitness of superordinate proposition could influence learners' construction of mental representations, the effects of superordinate inference generation on text comprehension remained unclear. To address these issues, Experiment 5 investigated the effects of task instructions on superordinate inference generation and text comprehension using superordinate inference and written recall tasks.

Thus, Experiment 5 was conducted in order to examine whether task instructions facilitate superordinate inference generation among Japanese EFL learners (RQ5-1) and whether task-induced strategic processing affects text comprehension (RQ5-2). In particular, Experiment 5 compared the difference between the Task condition (i.e., reading expository texts in order to understand the message conveyed by the writer) and the Control condition (i.e., reading texts for comprehension) in the performance of superordinate inference and written

recall tasks.

The results of the superordinate inference task showed that the success of superordinate inference was unrelated to task instructions. However, it was related to participants' L2 reading proficiency. Specifically, learners were likely to construct narrow representations based on each paragraph's main idea rather than the overall text. This suggested that Japanese EFL learners had difficulty integrating and constructing information distributed across paragraphs, regardless of task instructions. Furthermore, task instruction effects were not found in the written recall task. Although the present study aimed to induce learners to generate superordinate inferences, the task instructions might have been somewhat ambiguous and less strategic to facilitate learners' reading of expository texts. Therefore, the effects of task instructions on superordinate inference generation and text comprehension was further examined in Experiment 6.

The aim of Experiment 6 was to examine the following three points: effects of the reading task that induced EFL learners to integrate information distributed among paragraphs on superordinate inference generation (RQ6-1), cognitive processes during reading (RQ6-2), and text comprehension after reading (RQ6-3). Experiment 6 gave the integration task where the participants were required to integrate information across paragraphs after reading each paragraph. In this experiment, participants completed think-aloud, superordinate inference, and written recall tasks.

The results of the superordinate inference task showed that the integration task given in the present study effectively facilitated superordinate inference generation. In Experiment 6, the reading task required readers to repeatedly access previous information that was stored in long-term memory. As a result, learners constructed globally coherent representation of the expository text more easily in the Task than Control condition. Furthermore, the results of the written recall task revealed that such the integration task effectively promoted readers' text comprehension.

In contrast, the results of the think-aloud task demonstrated that cognitive processing during reading did not differ according to reading conditions. In particular, data indicated that EFL learners devoted too many resources to lower-level processing in expository reading. Consequently, they could not allocate remaining resources to engage in higher-level conceptual processing in accordance with the reading task. However, further analysis demonstrated that the relationships between processes and products differed contingent upon reading conditions, indicating that learners engaged in each process depending on the particular task. These results suggested that the integration task did not directly change learners' processes during reading. Instead, the task influenced learners' reading goals, resulting in better text comprehension in the Task condition.

Based on the results obtained from Experiments 4 to 6, the following three points should be discussed, respectively: (a) whether superordinate inference generation is the result of automatic or strategic processing in expository text comprehension? automatic or strategic processing in expository reading comprehension, (b) the effects of superordinate inference generation on text comprehension, and (c) cognitive resource allocation in expository reading.

Is superordinate inference generation the result of automatic or strategic processing in expository text comprehension?

As stated in section 5.1.1, theories of text comprehension have inconsistent conclusions as to whether global inferences are automatically or strategically generated during reading (e.g., Graesser et al., 1994; McKoon & Ratcliff, 1992). Furthermore, compared to existing narrative comprehension studies, there are limited studies on inferences during reading of expository texts (Lorch, 2015 for review).

In Experiment 4, the results of the inference verification task showed that readers judged an implicit superordinate proposition as highly appropriate as an explicit superordinate

proposition. Although this result suggested that EFL learners were able to spontaneously generate superordinate inference, even when not instructed to do so, this type of inference generation should be replicated with another methodology. Thus, Experiment 5 used a superordinate inference task and demonstrated that EFL learners had difficulty generating superordinate inference, regardless of task instructions. Instead, they tended to construct narrow text representations based on the main idea conveyed in a single paragraph rather than the overall text. In Experiment 6, when participants were instructed to integrate information across paragraphs, superordinate inference was facilitated.

These results suggested that superordinate inference represents strategic rather than automatic processing in EFL reading comprehension. The present study's finding about strategic generation of superordinate inference is inconsistent with Ritchey (2011), who reported that generation of superordinate inference occurred in L1 reading comprehension, regardless of reading goals (i.e., reading for verification/summary). According to Morishima (2013), while both L1 and L2 readers were able to maintain local text coherence, constructing global coherence was more difficult for L2 readers than L1 readers due to limited resource allocation for discourse-level processes. Many other studies also reported that L2 readers allocated fewer cognitive resources to higher-level processing compared to L1 readers (Horiba, 1996, 2000). Considering the differences between L1 and L2 reading, it seems reasonable that, in the present study, superordinate inference generation occurred based on the specific reading goal.

Since the present study did not directly measure the on-line generation of inferences during the time course of reading, whether superordinate inference is automatically or strategically generated should be further examined. However, given that superordinate inference was rarely produced in participants' think-aloud data (Experiment 6) and that superordinate inference generation was facilitated by integration task (Experiment 6) but not

by less strategic instructions (Experiment 5), it can be concluded that superordinate inference was not a spontaneous, effortless, and routine process in L2 expository reading.

The effects of superordinate inference generation on text comprehension

Superordinate inference is a kind of global inference that contributes to building global text coherence; however, whether and how it affects text comprehension was not yet clear. Although few previous studies directly examined this point, some reported results with implications related to the present study. For example, previous L1 studies demonstrated that when a superordinate proposition was explicitly mentioned as a title or topic sentence, readers used it to understand the gist of a text (e.g., Lorch et al., 1993; Sanchez et al., 2001). Conversely, Ushiro et al. (2008) demonstrated that when the superordinate proposition was not explicitly stated in the text, there was a weaker connection between some information, leading to building less coherent mental representation of the L2 text.

The results of Study 2 were consistent with the findings of related literature in that understanding superordinate information facilitated readers' text comprehension in a top-down processing manner. Although Experiment 4 did not directly measure participants' text comprehension, the inference verification task suggested that mental representations of the expository text were influenced by the explicitness of the superordinate proposition. Specifically, when the superordinate proposition was implicit, Japanese EFL learners were unlikely to suppress the activation of information inconsistent with the overall passage. Experiments 5 and 6 further examined EFL learners' text comprehension by manipulating task instructions. Here, the results demonstrated that when learners were given the integration task, superordinate inference generation was facilitated and led to better expository text comprehension. Therefore, this study concluded that explicit or inferred superordinate propositions helped learners to construct more stable and elaborative text representations in

their memories, which led to better text comprehension.

Cognitive resource allocation in L2 expository reading

To test Japanese EFL learners' cognitive resource allocation in expository reading, task instructions were manipulated. In Experiment 5, both superordinate inference generation and text comprehension were influenced by L2 reading proficiency, but not by task instructions. Although Experiment 5 aimed to induce learners to generate superordinate inferences, the task instructions might have been less strategic for facilitating learners' engagement in specific processing. Subsequently, Experiment 6 gave the reading task aimed at integrating information across paragraphs. The results revealed that while text comprehension was facilitated by the task, reading processes did not differ between reading conditions. Thus, Study 2 suggested that Japanese EFL learners had difficulty controlling their processes during expository reading based upon the reading task. Generally, expository texts include more unfamiliar concepts and have more complex structure compared to narrative texts (see section 2.1.2). In fact, Shimizu (2015) demonstrated that additional lower-level processing was needed for L2 readers to comprehend an expository than narrative text. Therefore, in the present study, learners primarily engaged in language competence-based processing rather than task-based processing.

On the other hand, further analyses revealed that the task instructions affected the relationship between comprehension processes and products. In particular, it was suggested that, although the proportion of think-aloud comments was similar in the two conditions, the standards of coherence (e.g., van den Broek et al., 1995, 2001) might differ depending on the task instructions. Specifically, while Japanese EFL learners set their standards for simple text comprehension when they were not given any strategic instructions, they set standards for superordinate inference generations when task instruction were provided.

Figure 5.2 illustrates construction of global coherence of expository texts in Japanese

EFL learners' reading comprehension, focusing on the task instructions and L2 reading proficiency on learners' standards of coherence, cognitive processes during reading, and products of reading comprehension. First, the task instructions for constructing implicit superordinate propositions did not facilitate either superordinate inference generation or text comprehension. On the other hand, when learners engaged in the integration task, they set their standards of coherence in order to construct superordinate propositions of the text rather than simply comprehend the text itself. Although learners' cognitive processes during reading were not directly influenced by the task engagement, integration processes, which occurred when they engaged in the superordinate inference task, facilitated superordinate inference generation and construction of robust mental representations.

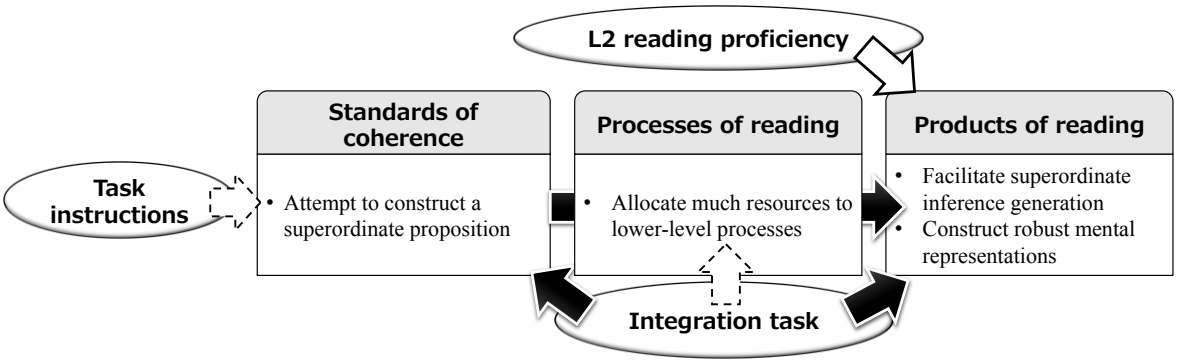

Figure 5.2. Construction of global coherence of expository texts in Japanese EFL learners' reading comprehension.

Construction of global coherence of expository texts in Japanese EFL learners' reading comprehension shown in Figure 5.2 can be explained by theories and models regarding reading comprehension and inference generation. The present study concluded that superordinate inference was not automatically generated during L2 reading, while it was strategically generated when task instructions were directed at specific reading goals. The present study supported the notion of the minimalist hypothesis, in that global inference is only generated

strategically and not automatically (McKoon & Ratcliff, 1992).

In addition, the effects of task instructions on inference generation and text comprehension can be accounted for by the landscape model (van den Broek et al., 1996). In the present study, while the task instructions aimed at superordinate inference generation did not affect the processes and products of reading comprehension, the learners' standards of coherence were altered by the integration task. Although the integration task did not directly change the cognitive processes during reading, it affected the relationships between the processes and the products. Specifically, when the learners were given the integration task, they generated inferences in order to achieve their goal (i.e., to perform the integration task) rather than to simply comprehend the text itself.

Furthermore, although the present study did not directly investigate the activation of individual concepts during reading, which is assumed in the landscape model, the results of the superordinate task suggested the possibility that the learners attempted to retrieve previous information and maintain coherence between previous and current information while engaging the integration task. Such processes may strengthen the activation levels of individual concepts, which leads to building the robust representations of expository texts.

5.3 Similarities and Differences in Constructing Globally Coherent Representations Between Narrative and Expository Reading

As previously noted, the present study focused on two different inference types: thematic inference and superordinate inference. Thus, constructions of global coherence of narrative and expository texts were examined in Study 1 and Study 2, respectively. Although this study did not directly compare narrative and expository reading, some similarities and differences between these two text types were suggested. The results of the six experiments are discussed from the following two perspectives: automaticity of global inference generation, and cognitive

resource allocation.

Automaticity of global inference generation

As reviewed in section 2.2.1, there is disagreement between the minimalist hypothesis and constructionist theory, especially in terms of the generation of inferences that establish global coherence. According to the minimalist hypothesis (McKoon & Ratcliff, 1992), inferences are automatically generated when they are based on quickly and easily available information that are necessary for local text coherence. Global inference requires readers to link elements that are being currently processed substantially earlier elements that are no longer retained in short-term working memory and/or to relevant world knowledge. Therefore, it is claimed that global inferences are strategically generated only when readers have specific goals and engage in strategic processing to achieve these goals. In contrast, constructionist theory asserts points that the reader attempts to construct a meaningful situation model that is coherent at both local and global levels (Graesser et al., 1994). Therefore, according to this theory, global inference can be generated during the course of reading comprehension.

Although the present study did not directly investigate the on-line generation of thematic and superordinate inference, it can be concluded that Japanese EFL learners' used strategic rather than automatic processing in text comprehension based on the following results derived from the six experiments. First, while learners could judge presented themes and superordinate propositions to be appropriate (i.e., the inference verification task in Experiments 1 and 4), they were unable to construct overall representations of the text when not provided with task instructions (i.e., the thematic/superordinate inference task in Experiments 2, 5, and 6). Second, participants' produced few comments on thematic and superordinate inferences in the think-aloud data (Experiments 3 and 6).

Some previous L1 and L2/EFL studies, have examined whether or not specific inference

types are automatically generated. For example, Ushiro et al. (2012) revealed that Japanese EFL learners activated causal bridging inferences on-line, while predictive inference generation was somewhat delayed. Additionally, Nahatame's (2014) study focused on predictive inference and demonstrated that learners automatically generated predictive inferences during reading only when provided with strategy instructions. The current research offers new insight into inference generation in L2/EFL reading. Thus far, only a few studies have focused on what types of inferences are generated during L2 reading comprehension so far. Therefore, generation of specific inference types, including global coherence inferences, should be empirically investigated in future research.

Cognitive resource allocation in narrative and expository reading

Many previous studies have examined whether task instructions given before reading affect reading processes and post-reading text comprehension. As reviewed in section 2.3, a substantial number of studies provided evidence that L1 readers are able to adjust their cognitive processes and strategies in accordance with their reading goals (e.g., Bråten & Strømsø, 2009; Linderholm & van den Broek, 2002; Magliano et al., 1999; Narvaez et al., 1999; van den Broek et al., 2001). Conversely, cognitive resource allocation during reading is likely to be more complex and unstable for L2 than L1 reading (e.g., Horiba, 2000, 2013; Yoshida, 2012).

In Experiments 2 and 5, participants were told that they would complete a thematic/superordinate inference task after reading; therefore, they needed to think about what the writer intended to convey throughout the overall text. The results showed that while task instructions facilitated thematic inference generation and text comprehension in narrative reading, they did not affect superordinate inference generation and text comprehension in expository reading. Thus, task instructions where the learners were only instructed to think about the writer's message were not sufficient to alter learners' reading goals in expository

reading. However, Experiment 6 gave the reading task to engage in integrating information across paragraphs and demonstrated that these instructions facilitated superordinate inference generation and text comprehension in expository reading.

As for cognitive processes during reading, the think-aloud data from Experiment 3 showed that Japanese EFL learners' cognitive resource allocation was partially altered by task instructions. Specifically, task instructions reduced the allocation of cognitive resources to lower-level linguistic processing (e.g., analysis of each word and sentence). Conversely, Experiment 6 demonstrated that cognitive processes did not change dependent on the reading task.

According to the results of Study 1 (Experiments 1 to 3) and Study 2 (Experiments 4 to 6), similarities and differences between narrative and expository reading among Japanese EFL learners can be summarized in the following three points.

First, the effects of task instructions were likely to appear more clearly in narrative than expository reading. While the narrative stories used in Study 1 consisted of one paragraph, Study 2's expository texts consisted of four paragraphs. Therefore, learners were required to integrate more globally distributed information to build coherent mental representations of expository than narrative texts. As the result of limited resource allocation for discourse-level processing, L2/EFL learners had more difficulty relating earlier and present information during reading compared to L1 readers (Morishima, 2013). Consequently, the effects of task instructions were only found in expository reading when the instructions induced learners to strategically relate globally distributed text information.

Second, while expository reading was more likely to be influenced by learners' L2 reading proficiency than task instructions, narrative reading was more likely to be influenced by task instructions than L2 reading proficiency. As reviewed in 2.1.2, expository reading was more difficult than narrative reading because of less proficiency and prior knowledge as well

as greater complexity in expository texts (Freedle & Halle, 1979; Wolfe, 2005; Wolfe & Mienko, 2007; Wolfe & Woodwyk, 2010). Shimizu (2015) also demonstrated that lower proficiency L2 learners tended to allocate more cognitive resources to lower-level processing in expository than narrative reading. Regarding the difficulty and complexity of expository reading suggested in previous studies, it is reasonable that expository reading task effects were relatively small because learners primarily engaged in language competence-based processing rather than task-based processing.

Finally, although cognitive resource allocation was limited and unstable in EFL reading comprehension, the results suggested that learners tried to alter cognitive processes according to reading goals provided in task instructions for both narrative and expository reading. Additionally, Experiment 3's questionnaire indicated that Japanese EFL learners' attention shifted from understanding individual pieces of information to connecting text elements to build globally coherent representations. Although Experiment 6 did not utilize such a questionnaire, the results about the relationships between think-aloud comments and the written recall task suggested that learners engaged in specific processing during reading in order to achieve the goals induced by task instructions.

Effects of linguistic difficulties of experimental passages used in the present study

Although some similarities and differences in constructing globally coherent representations between narrative and expository reading were discussed above, one possible argument against the present study might be that experimental passages used in Study 1 and Study 2 were different in terms of linguistic difficulty (e.g., vocabulary, syntactic structure, text length) as well as text type. The present study did not control these factors completely because it did not aim to directly compare the two text types. Therefore, to interpret the results appropriately, it seems necessary to interpret the obtained results not only from the text type but

also the difficulty of the texts themselves.

As the narrative passages used in Study 1 were adapted from Seifert et al. (1984) which targeted L1 readers, they were modified for Japanese EFL learners. As a result, lower frequency words were changed into higher frequency words, and unfamiliar phrases and complex sentences were simplified. Moreover, each passage consisted of only one paragraph and the number of words in each passage was around 80 (see Table 3.2). In contrast, experimental passages used in Study 2 were adopted from the reading section of the Second Grade STEP test. Regarding the proficiency level of participants in the present studies, these materials were not too difficult for them; however, the levels of vocabulary, syntactic structure, and text length may have made it more difficult for the participants to read them.

Although the present study did not collect the data of the difficulty and readability of the experimental passages for the participants, some data showed the possibility that passages used in Study 2 were more difficult than those in Study 1. For example, while the mean recall production rates in narrative texts were around 40-50% (see Tables 3.12, 3.19), those in expository texts were 25-35%. Furthermore, the think-aloud data showed that the mean proportion of comments on word and sentence analyses was around 55-80% in narrative texts (see Table 3.27), whereas that of expository texts was more than 80% (see Table 4.16). These results suggest that it was not only the text type but also linguistic difficulty that made it difficult to construct globally coherent mental representations of the texts.

When the text difficulty increases, L2 learners have difficulty maintaining prior elements in the text, which leads to less inference generation (Barry & Lazarte, 1998). As a result, the effects of task instructions were smaller in expository reading than narrative reading. Yoshida (2012) also explained that the task instructions did not have positive effects on L2 text comprehension because "linguistic constraints may have overridden the task effects (p. 19)." The results of the present study were consistent with these explanations. Therefore, it should be

concluded that linguistic difficulty of the texts, as well as text types themselves, may have attributed to the varying patterns of globally coherent mental representations of text comprehension.

Chapter 6
Conclusion

6.1 Summary of Findings

The present study aimed to investigate how Japanese EFL learners construct globally coherent mental representations of texts, focusing on thematic inference in narrative reading and superordinate inference in expository reading. The main findings of the present study can be summarized as follows:

In narrative reading, although Japanese EFL learners had difficulty comprehending implicit themes of narrative texts through inference generation, the task instructions aimed at thematic inference facilitated strategic processing and it contributed to successful thematic inference generation and better text comprehension. When the task instructions were given, EFL learners' reading goals shifted to building global coherence of texts, understanding causal relationships between textual information, and understanding characters' intentions. However, the effects of task instructions only appeared in parts of the cognitive processes during reading. Moreover, the effects of L2 reading proficiency were found in text comprehension and the cognitive processes.

In expository reading, Japanese EFL learners had difficulty comprehending implicit superordinate propositions of expository texts through inference generation. The task instructions aimed at superordinate inference did not affect inference generation and text comprehension, whereas the integration task helped learners to construct superordinate propositions and better comprehend the text. The integration task did not directly affect cognitive processes during reading; however, it facilitated learners' retrieval and integration processes while they engaged the task. Moreover, the effects of L2 reading proficiency were found to facilitate success in superordinate inference generation and text comprehension.

6.2 Limitations of the Present Studies and Suggestion for Future Research

Although the present study may provide new insights into L2/EFL reading comprehension, it is necessary to recognize some limitations. In order to pursue my research interests, the following three limitations should be further considered: (a) the present study's methodology, (b) participants' characteristics, and (c) direct comparisons of narrative and expository reading.

First, although the present study selected the most appropriate methodologies to investigate each research question, some elements should be reconsidered and improved. For example, in order to observe Japanese EFL learners' inference generation, the inference verification task (Experiments 1 and 4) and the thematic/superordinate inference tasks (Experiments 2, 5, and 6) were adopted. These tasks are valid measures of deeper, situational understanding of texts and have been widely used in previous studies to test the ability to make inferences (Bråten & Strømsø, 2009; Campion & Rossi, 2001; Kurtz & Schober, 2001; Rapp & Gerrig, 2006; Royer et al., 1996). However, they do not allow for investigation of on-line inference generation. For instance, previous studies have examined the generation of other inference types, especially local inference, using other methodologies such as a lexical decision task, a recognition task, a meaningfulness judgment task, and eye tracking. Moreover, the available methodology for investigating global inferences may be more limited than for local inferences because the superordinate inference was triggered by lengthy text passages rather than single words, clauses, and sentences. However, combining an on-line methodology with the off-line tasks used in the present study will provide more objective and precise evidence for global inference generation in L2/EFL reading comprehension.

In addition, in the present study, the think-aloud task used in Experiments 3 and 6 is the most suitable methodology for examining what cognitive processes learners engaged in during reading; however, the limitations of the method should be noted (Bowles, 2010; Israel, 2015

for review). For example, the think-aloud protocol does not reflect all processes or strategies used by L2 readers in text comprehension. Thus, other on-line methods such as reading time and eye movements, which reflect readers' cognitive resource allocation during reading, can be combined with the think-aloud method.

Second, the present study should be replicated with a larger sample of participants with more varied levels of L2 reading proficiency and various types of task instructions. Inconsistent with previous L1 studies (Bohn-Gettler & Kendeou, 2014; Linderholm & van den Broek, 2002), the present study did not find a clear interaction between task instructions and reader-related factors. It should also be noted that all participants in the six experiments were students from the same university; therefore, there may have been less variability in their L2 reading proficiency level. If participants' proficiency level and age had been more varied, the interaction between task instructions and L2 reading proficiency could have been observed more clearly.

Furthermore, future research should more focus on the characteristics of reading tasks. In the present research, Study 1 focused on the effects of task instructions aimed at thematic inference while Study 2 focused on the effects of task instructions aimed at superordinate inference and the integration task. As these studies only compared an experimental condition and a control condition, it was still unclear as to which type of task was most effective for EFL learners and what kind of learners benefitted from reading tasks. Compared to L1 reading, the effects of reading tasks on L2 reading comprehension were unstable and complex (Horiba, 2000, 2013; Yoshida, 2012). Therefore, more research on comparison of various types of reading by controlling the cognitive demands of reading tasks is required.

Third and finally, the present study did not conduct direct comparisons between narrative and expository reading comprehension among Japanese EFL learners. Although reading comprehension of the two text types were examined respectively, some differences between these text types were found. However, as reported in Horiba (2000, Experiment 1) and Shimizu

(2015), experiments that directly compare narrative and expository texts could more clearly demonstrate the characteristics of L2/EFL reading comprehension. Moreover, as stated in general discussion, the experimental passages used in Study 1 and Study 2 may differ in terms of linguistic difficulty as well as text type. To make the influence of text type clearer, future research should control linguistic factors such as vocabulary, syntactic structure, and text length as much as possible while directly comparing two text types. Such comparisons will allow us to examine more detailed processes regarding construction of globally coherent mental representations of texts among Japanese EFL learners.

6.3 Pedagogical Implications

The present study concluded that both thematic and superordinate inferences involve strategic rather than automatic processing in Japanese EFL learners' text comprehension. In practical English classes, although learners can translate each sentence into Japanese, learners frequently cannot understand the meaning of the whole text. Specifically, without any instructions, L2 or EFL learners sometimes tend to engage in comprehending each word or sentence. Consequently, they fail to understand the overall meaning of the text. Since reading comprehension entails visual communication between writers and readers, discourse comprehension requires readers to understand the writer's message by constructing both globally and locally coherent mental representations of texts. Therefore, learners require appropriate educational interventions. The present study's findings have both educational and theoretical implications. The pedagogical implications are summarized from the following three perspectives: reading materials for improving construction of globally coherent mental representations, effective instructions for constructing globally coherent mental representations, and cognitive resource allocation during EFL reading.

Reading materials for improving construction of globally coherent mental representations

In order to improve EFL learners' construction of globally coherent mental representations, teachers need to select appropriate reading materials in terms of explicitness and comprehensibility of authors' messages (themes, superordinate propositions). As for the explicitness of author's messages, Experiment 1 and Experiment 4 suggested that Japanese EFL learners at college student level had difficulty understanding authors' messages when they were not explicitly stated in the texts. Therefore, in order for learners to go beyond the literal level of comprehension and construct authors' messages based on context, reading materials where the theme or superordinate proposition are not explicitly mentioned should be prepared for the learners.

As for the comprehensibility of author's messages, readers do not attempt globally coherent representations of the text when the text lacks global coherence and a message (Graesser et al., 1994). Therefore, although teachers do not necessarily need to select the materials from which only one theme is derived, they should prepare materials where most learners can find common messages.

For example, as training for narrative theme comprehension, *fables* can be used in reading instruction. Fables are a kind of narrative where readers define particular moral points based on the characters' actions and outcomes (Dorfman & Brewer, 1994). Well-known stories such as "The Boy Who Cried Wolf" and "The Dog and The Shadow" are typical examples for fables. By adopting fables as materials in reading instruction, students focus not only on local and explicit information in the text but also author's messages conveyed through the overall passage.

In the case of expository reading, teachers should consider what types of text structure can improve construction of globally coherent mental representations. As reviewed in section 2.1.2, the text structure of expository texts varies more than that of narrative texts (e.g., Meyer et al., 1980; Meyer & Freedle, 1984). Considering the characteristics of each text structure, the

expository texts with problem/solution or comparison structure seem more appropriate for reading instruction focusing on superordinate inference generation than those with causation or collection/description. The reason for this is that these types of passages were better organized and the writers are likely to convey their intended message as the solution to the problems or as the superior idea.

Effective instructions for constructing globally coherent mental representations in EFL reading

The present study demonstrated the effectiveness of task instructions (Experiments 2, 3) and integration tasks (Experiment 6) on constructing globally coherent mental representations in EFL reading. Based on the results, the present study suggests effective pre-reading, mid-reading, and post-reading instructions as follows.

When a teacher gives the pre-reading instruction, "read the text in order to understand the author's message" to students, their standards of coherence are likely to be set to construct global mental representations and they will strategically read the text in a top-down processing manner in order to achieve the goal. This pre-reading instruction will be adequate when students read a relatively short and simple passage, similar to passage used in Study 1. In contrast, when students read a complex text as in Study 2, more direct intervention may be needed during reading.

As a mid-reading task, the integration task adopted in Experiment 6 will be effective especially when students have difficulty integrating information across paragraphs. For example, when a teacher tells students to state the possible message conveyed through the overall text after reading each paragraph, students' retrieval and integrative processes will be facilitated. In this task, teachers should be careful that students focus on integrating multiple paragraphs rather than simply summarizing the current paragraph.

As a post-reading activity, a teacher should ask, "what is the message of this passage?"

again and allow the students to discuss it in small groups. As it is highly likely that students have constructed various messages for each text, group work can be one of the effective methods. In discussions, students give various answers, some of which are inappropriate for the text. When this occurs, rather than immediately pointing out that the answer is incorrect, the teacher should ask students to provide reasons for their answers. This will reveal the source of the students' difficulty (e.g., failure in comprehending the text itself, selecting important elements, and integrating information).

In order to make the most of task instructions and reading tasks, teachers should also consider the characteristics of readers, texts, and instructions as well as the interactions between them. For example, Horiba (2000, 2013) claimed that task instruction effects will be clearer when task instructions are specific because this forces learners to alter particular processing behaviors while reading or engaging in post-reading tasks. Moreover, EFL learners cannot take advantage of task instructions related to reading comprehension when task instructions and participants' proficiency levels do not match (Yoshida, 2012).

Cognitive resource allocation during EFL reading

Consistent with previous studies, the present study demonstrated that the processes of EFL learners were unstable and inflexible compared to L1 readers, resulting from an overload on lower-level linguistic processing (Horiba, 1996, 2000; Morishima, 2013; Shimizu, 2015). That is, Japanese EFL learners engaged in excessive lower-level processing, which led to less resource allocation to higher-level processing during reading.

To overcome this problem, concerning the one-shot instruction, the present study suggested that the pre-reading task instructions aimed at global inferences can prevent learners from allocating too many cognitive resources to lower-level linguistic processing. Conversely, the effects of task instructions on cognitive resource allocation in reading comprehension can

be temporary. Therefore, teachers need to train students to flexibly control their reading processes using long-term instructions. Horiba (2013) claims that carefully designed tasks allow students to experience different processing modes and develop reading strategies. For example, teachers could explicitly teach higher-level conceptual strategies such as knowledge-based inference generation using text structure and monitoring according to the given reading tasks. It would be desirable that students be able to select the appropriate reading strategy according to the tasks by themselves.

In addition, teachers should note that the way students read texts in the classroom setting can influence their standards of coherence. Specifically, if a teacher pays too much attention to explaining each word and syntactic structures in the text or asking only fact-finding questions, students will settle for lower standards of coherence and will not engage in strategic processes for coherence building. Therefore, to encourage active and deeper-level processing in learners, teachers should ask high-level questions as well as lower-level questions, and give specific reading goals to achieve rather than simply asking that students comprehend the text itself.

6.4 Concluding Remarks

In L2 or EFL reading, teachers sometimes believe that their students have successfully read a passage when they can perfectly translate a text. However, given that reading is an activity in which readers and writers interact to construct meaning, a final goal of learners' reading comprehension is to understand the writer's meaning throughout the overall text rather than only at the surface level or in parts of the text. Although relatively longer texts are frequently used in L2 or EFL reading education, whether readers grasp the overall text's image has not been well considered or empirically examined. Consequently, the current study examined EFL learners' globally coherent mental representations of the texts, focusing on thematic inference in narrative reading and superordinate inference in expository reading.

Although, as stated above, the current study has several limitations and controversial elements, it can provide new insight that clarifies aspects of reading comprehension among Japanese EFL learners. The present study is significant in that it demonstrates not only whether readers can or cannot generate global inference, but also how to solve difficulties that occur in EFL learners' reading processes. However, to obtain more theoretical and pedagogical implications for L2/EFL reading, further research investigating reading comprehension at a global level is needed.

References

Barry, S., & Lazarte, A. A. (1998). Evidence for mental models: How do prior knowledge, syntactic complexity, and reading topic affect inference generation in a recall task for nonnative readers of Spanish? *The Modern Language Journal, 82,* 176–193. doi:10.1111/j.1540-4781.1998.tb01191.x

Bohn-Gettler, C. M., & Kendeou, P. (2014). The interplay of reader goals, working memory, and text structure during reading. *Contemporary Educational Psychology, 39,* 206–219. doi:10.1016/j.cedpsych.2014.05.003

Bowles, M. A. (2010). *The think-aloud controversy in second language research*. New York, NY: Routledge.

Bråten, I., & Strømsø, H. I. (2009). Effects of task instruction and personal epistemology on the understanding of multiple texts about climate change. *Discourse Processes, 47,* 1–31. doi:10.1080/01638530902959646

Brown, A. L., & Day, J. D. (1983). Macrorules for summarizing texts: The development of expertise. *Journal of Verbal Learning and Verbal Behavior, 22,* 1–14. doi:10.1016/S0022-5371(83)80002-4

Brown, A. L., Day, J. D., & Jones, R. S. (1983). The development of plans for summarizing texts. *Child Development, 4,* 968–979. doi:10.2307/1129901

Campion, N., & Rossi, J. P. (2001). Associative and causal constraints in the process of generating predictive inferences. *Discourse Processes, 31,* 263–291. doi:10.1207/S15326950dp31-3_3

Carrell, P. L. (1984). The effects of rhetorical organization on ESL readers. *TESOL Quarterly, 18,* 441–469. doi:10.2307/3586714

Carrell, P. L. (1992). Awareness of text structure: Effects on recall. *Language Learning*, *42*, 1–20. doi:10.1111/j.1467-1770.1992.tb00698.x

Coté, N., Goldman, S. R., & Saul, E. U. (1998). Students making sense of informational text: Relations between processing and representation. *Discourse Processes*, *25*, 1–53. doi:10.1080/01638539809545019

Dorfman, M. H., & Brewer, W. F. (1994). Understanding the points of fables. *Discourse Processes*, *17*, 105–129. doi:10.1080/01638539409544861

Freedle, R., & Halle, G. (1979). Acquisition of new comprehension schemata for expository prose by transfer of a narrative schema. In R. Freedle (Ed.), *New directions in discourse processing* (pp. 121–135). Norwood, NJ: Ablex.

Geiger, J. F., & Millis, K. K. (2004). Assessing the impact of reading goals and text structures on comprehension. *Reading Psychology*, *25*, 93–110. doi:10.1080/02702710490435637

Gernsbacher, M. A. (1990). *Language comprehension as structure building.* Hillsdale, NJ: Erlbaum.

Gernsbacher, M. A. (1996). The structure-building framework: What it is, what it might also be, and why. In B. K. Britton & A. C. Graesser (Eds.), *Models of understanding text* (pp. 289–311). Mahwah, NJ: Erlbaum.

Gernsbacher, M. A., Varner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 430–445.

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

Grabe, W., & Stoller, F. L. (2002). *Teaching and researching in reading*. Harlow, UK: Pearson Education.

Graesser, A. C., & Goodman, S. M. (1985). Implicit knowledge, question answering, and the representation of expository text. In B. K. Britton & J. B. Black (Eds.), *Understanding expository text* (pp. 109–171). Hillsdale, NJ: Lawrence Erlbaum Associations.

Graesser, A. C., Pomeroy, V. J., & Craig, S. D. (2002). Psychological and computational research on theme comprehension. In M. Louwerse & W. van Peer (Eds.), *Thematics: Interdisciplinary studies* (pp. 19–34). Amsterdam, Netherlands: John Benjamins.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101*, 371–395. doi:10.1037/0033-295X.101.3.371

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London, UK: Longman.

Horiba, Y. (1996). Comprehension processes in L2 reading: Language competence, textual coherence, and inferences. *Studies in Second Language Acquisition, 18*, 433–473. doi:10.1017/S0272263100015370

Horiba, Y. (2000). Reader control in reading: Effects of language competence, text type, and task. *Discourse Processes, 29*, 223–267. doi:10.1207/S15326950dp2903_3

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

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

Horiba, Y., van den Broek, P., & Fletcher, C. R. (1993). Second language readers' memory for narrative texts: Evidence for structure preserving top-down processing. *Language Learning, 43*, 345–372. doi:10.1111/j.1467-1770.1993.tb00618.x

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

Israel, S. E. (2015). *Verbal protocols in literacy research: Nature of global reading development*. New York, NY: Routledge.

JACET. (2003). *JACET list of 8000 basic words*. Tokyo, Japan: Author.

Johnson-Laird, P. N. (1983). *Mental model*. Cambridge: Harvard University Press.

Kaakinen, J. K., & Hyönä, J. (2005). Perspective effects on expository text comprehension: Evidence from think-aloud protocols, eyetracking, and recall. *Discourse Processes*, *40*, 239–257. doi:10.1207/s15326950dp4003_4

Kaakinen, J. K., Hyönä, J., & Keenan, J. M. (2002). Perspective effects on online text processing. *Discourse Processes*, *33*, 159–173. doi:10.1207/S15326950DP3302_03

Kendeou, P., van den Broek, P., Helder, A., & Karlsson, J. (2014). A cognitive view of reading comprehension: Implications for reading difficulties. *Learning Disabilities Research & Practice*, *29*, 10–16. doi: 10.1111/ldrp.12025

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*, 163–182. doi:10.1037/0033-295X.95.2.163

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*, 363–394.

Kintsch, W., Welsch, D., Schmalhofer, F., & Zimny, S. (1990). Sentence memory: A theoretical analysis. *Journal of Memory and Language*, *29*, 133–159. doi:10.1016/0749-596X(90)90069-C

Kintsch, W., & Young, S. R. (1984). Selective recall of decision-relevant information from texts. *Memory & Cognition*, *12*, 112–117. doi:10.3758/BF03198424

Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. New York, NY: Cambridge University Press.

Kurtz, V., & Schober, M. F. (2001). Readers' varying interpretations of theme in short fiction. *Poetics, 29,* 139–166. doi:10.1016/S0304-422X(01)00040-7

Lee, J. F. (1986). On the use of the recall task to measure L2 reading comprehension. *Studies in Second Language Acquisition, 8*, 201–212.

Lehr, S. (1988). The child's developing sense of theme as a response to literature. *Reading Research Quarterly, 23,* 337–357. doi:10.2307/748046

Linderholm, T., & van den Broek, P. (2002). The effects of reading purpose and working memory capacity on the processing of expository text. *Journal of Educational Psychology, 94,* 778–784. doi:10.1037/0022-0663.94.4.778

Long, D. L., Oppy, B. J., & Seely, M. R. (1994). Individual differences in the time course of inferential processing. *Journal of Experimental Psychology: Learning, Memory and Cognition, 20,* 1456–1470.

Long, D. L., Oppy, B. J., & Seely, M. R. (1997). Individual differences in readers' sentence- and text-level representations. *Journal of Memory and Language, 36,* 129–145. doi:10.1006/jmla.1996.2485

Lorch, R. F. Jr., (2015). What about expository text? In E. J. O'Brien, A. E. Cook, & R. F. Lorch Jr. (Eds), *Inferences during reading* (pp. 348–361). Cambridge University Press.

Lorch, R. F. Jr., Lorch, E. P., & Inman, W. E. (1993). Effects of signaling topic structure on text recall. *Journal of Educational Psychology, 85*, 281–290.

Lorch, R. F. Jr., & van den Broek, P. (1997). Understanding reading comprehension: Current and future contributions of cognitive science. *Contemporary Educational Psychology, 22,* 213–246.

Magliano, J. P., Trabasso, T., & Graesser, A. C. (1999). Strategic processing during comprehension. *Journal of Educational Psychology, 91,* 615–629. doi:10.1037/0022-0663.91.4.615

Mandler, J. M. (1984). *Stories, scripts, and science: Aspects pf schema theory*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Mandler, J. M. (1987). On the psychological reality of story structure. *Discourse Processes*, *10*, 1–29. doi:10.1080/01638538709544656

Mandler, J. M., & Johnson, N. S. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology*, *9*, 111–151. doi:10.1016/0010-0285(77)90006-8

McCarthy, K. S., & Goldman, S. R. (2015). Comprehension of short stories: Effects of task instructions on literary interpretation. *Discourse Processes*, *52*, 1–24. doi:10.1080/0163853X.2014.967610

McMaster, K. L., Espin, C. A., & Broek, P. (2014). Making connections: Linking cognitive psychology and intervention research to improve comprehension of struggling readers. *Learning Disabilities Research & Practice*, *29*, 17–24. doi: 10.1111/ldrp.12026

McCrudden, M. T., Magliano, J. P., & Schraw, G. (2010). Exploring how relevance instructions affect personal reading intentions, reading goals, and text processing: A mixed methods study. *Contemporary Educational Psychology*, *35*, 229–241. doi:10.1016/j.cedpsych.2009. 12.001

McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review*, *99*, 440–466. doi:10.1037/0033-295X.99.3.440

Meyer, B. J., Brandt, D. M., & Bluth, G. J. (1980). Use of top-level structure in text: Key for reading comprehension of ninth-grade students. *Reading Research Quarterly*, *16*, 72–103. doi:10.2307/747349

Meyer, B. J., & Freedle, R. O. (1984). Effects of discourse type on recall. *American Educational Research Journal*, *21*, 121–143. doi:10.3102/00028312021001121

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

Mulder, G., & Sanders, T. J. (2012). Causal coherence relations and levels of discourse representation. *Discourse Processes*, *49*, 501–522. doi:10.1080/0163853X.2012.692655

Muramoto, T. (2000). Daini gengo no bunshou rikai katei ni oyobosu shujukudo no eikyo [The effects of second-language proficiency on text comprehension]. *Science of Reading*, *44*, 43–50.

Nahatame, S. (2014). Strategic processing and predictive inference generation in L2 reading. *Reading in a Foreign Language*, *26*, 54–77.

Narvaez, D. (1998). The influence of moral schemas on the reconstruction of moral narratives in eighth graders and college students. *Journal of Educational Psychology*, *90*, 13–24. doi:10.1037/0022-0663.90.1.13

Narvaez, D. (1999). Using discourse processing methods to study moral thinking. *Educational Psychology Review*, *11*, 377–393. doi:10.1023/A:1022061417089

Narvaez, D., van den Broek, P., & Ruiz, A. B. (1999). The influence of reading purpose on inference generation and comprehension in reading. *Journal of Educational Psychology*, *91*, 488–496. doi:10.1037/0022-0663.91.3.488

Noordman, L. G., Vonk, W., & Kempff, H. J. (1992). Causal inferences during the reading of expository texts. *Journal of Memory and Language*, *31*, 573–590. doi:10.1016/0749-596X(92)90029-W

Nuttall, C. (2005). *Teaching reading skills in a foreign language.* Oxford, UK: Macmillan Education.

Poynor, D. V., & Morris, R. K. (2003). Inferred goals in narratives: Evidence from self-paced reading, recall, and eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 3–9. doi:10.1037//0278-7393.29.1.3

Rapp, D. N., & Gerrig, R. J. (2006). Predilections for narrative outcomes: The impact of story contexts and reader preferences. *Journal of Memory and Language, 54,* 54–67. doi:10.1016/j.jml.2005.04.003

Rinck, M., & Weber, U. (2003). Who when where: An experimental test of the event-indexing model. *Memory & Cognition, 31,* 1284–1292. doi:10.3758/BF03195811

Ritchey, K. A. (2011). How generalization inferences are constructed in expository text comprehension. *Contemporary Educational Psychology, 36,* 280–288. doi:10.1016/j.cedpsych.2011.03.002

Royer, J. M., Carlo, M. S., Dufresne, R., & Mestre, J. (1996). The assessment of levels of domain expertise while reading. *Cognition and Instruction, 14,* 373–408. doi:10.1207/s1532690xci1403_4

Sanchez, R. P., Lorch, E. P., & Lorch, R. F. Jr. (2001). Effects of headings on text processing strategies. *Contemporary Educational Psychology, 26,* 418–428. doi:10.1006/ceps.2000.1056

Seifert, C. M., McKoon, G., Abelson, R. P., & Ratcliff, R. (1986). Memory connections between thematically similar episodes. *Journal of Experimental Psychology, 12,* 220–231. doi:10.1037/0278-7393.12.2.220

Shimizu, H. (2015). Generation of local and global bridging inferences in L2 reading comprehension. *JACET Journal, 59,* 75–92.

Singer, M. (1994). *Discourse inference processes*. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 479–516). San Diego: Academic Press.

Singer, M., Graesser, A. C., & Trabasso, T. (1994). Minimal or global inference during reading. *Journal of Memory and Language, 33,* 421–441. doi:10.1006/jmla.1994.1020

Singer, M., Harkness, D., & Stewart, S. T. (1997). Constructing inferences in expository text comprehension. *Discourse Processes, 24,* 199–228. doi:10.1080/01638539709545013

Singer, M., & Kintsch, W. (2001). Text retrieval: A theoretical exploration. *Discourse Processes*, *31*, 27–59. doi:10.1207/S15326950dp3101_2

STEP: The Society for Testing English Proficiency. (1997). *The EIKEN test in practical English proficiency grade 2*. Tokyo, Japan: Dai-X Shuppan.

STEP: The Society for Testing English Proficiency. (1997). *The EIKEN test in practical English proficiency grade pre 1*. Tokyo, Japan: Dai-X Shuppan.

STEP: The Society for Testing English Proficiency. (2009). *The EIKEN test in practical English proficiency grade pre 2*. Tokyo, Japan: Obunsha.

Tapiero, I. (2007). *Situation models and levels of coherence.* Mahwah, NJ: Lawrence Erlbaum Associates.

Thorndyke, P. W. (1977). Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, *9*, 77–110. doi:10.1016/0010-0285(77)90005-6

Till, R. E., Mross, E. F., & Kintsch, W. (1988). Time course of priming for associate and inference words in a discourse context. *Memory & Cognition*, *16*, 283–298. doi: 10.3758/BF03197039

Trabasso, T., & Sperry, L. (1985). Causal relatedness and importance of story events. *Journal of Memory and Language*, *24*, 595–611. doi:0749596X85900488

Trabasso, T., & van den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language*, *24*, 612–630. doi:0749596X8590049X

Trabasso, T., van den Broek, P., & Suh, S. (1989). Logical necessity and transitivity of causal relations in stories. *Discourse Processes*, *12*, 1–25. doi:10.1080/01638538909544717

Trabasso, T., & Wiley, J. (2005). Goal plans of action and inferences during comprehension of narratives. *Discourse Processes*, *39*, 129–162. doi:10.1080/0163853X.2005.9651677

Ushiro, Y., Nahatame, S., Hasegawa, Y., Kimura, Y., Hamada, A., & Tanaka, N. (2014). Narrative characters' goals and EFL readers' text comprehension: Focusing on goal explicitness. *ARELE: Annual Review of English Language Education in Japan*, *25*, 1–16.

Ushiro, Y., Nahatame, S., Hasegawa, Y., Shimizu, H., Kimura, Y., Tanaka, Y., & Nakagawa, C. (2012). Activation and encoding of bridging and predictive inferences in EFL reading comprehension. *JACET Journal*, *54*, 33–52.

Ushiro, Y., Nakagawa, C., Kai, A., Watanabe, F., & Shimizu, H. (2008). Construction of a macroproposition from supporting details: Investigation from Japanese EFL reader's summary and importance rating. *JACET Journal*, *47*, 111–125.

Ushiro, Y., Shimizu, H., Kai, A., Nakagawa, C., Takaki, S., Kobayashi, M., Satake, N., & Takano, D. (2010). Effects of causal networks on on-line and off-line narrative comprehension among Japanese EFL readers. *JACET Journal*, *51*, 39–53.

van den Broek, P. (1994). Comprehension and memory of narrative texts: Inference and coherence. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 539–588). San Diego: Academic Press.

van den Broek, P., Bohn-Gettler, C., Kendeou, P., Carlson, S., & White, M. J. (2011). When a reader meets a text: The role of standards of coherence in reading comprehension. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (pp.123–140). Charlotte, NC: Information Age Publishing.

van den Broek, P., Fletcher, C. R., & Risden, K. (1993). Investigations of inferential processes in reading: A theoretical and methodological integration. *Discourse Processes*, *16*, 169–180. doi:10.1080/01638539309544835

van den Broek, P., Lorch, R. F. Jr., Linderholm, T., & Gustafson, M. (2001). The effects of readers' goals on inference generation and memory for texts. *Memory & Cognition*, *29*, 1081–1087. doi:10.3758/BF03206376

van den Broek, P., Risden, K., Fletcher, C. R., & Thurlow, R. (1996). A 'landscape' view of reading: Fluctuating patterns of activation and the construction of a stable memory representation. In B. K. Britton & A. C. Graesser (Eds.), *Models of understanding text* (pp.165–187). Mahwah, NJ: Erlbaum.

van den Broek, P., Risden, K., & Husebye-Hartmann, E. (1995). The role of readers' standards for coherence in the generation of inferences during reading. In R. F. Lorch Jr. & E. J. O'Brien (Eds.), *Sources of coherence in reading* (pp. 353–373). Hillsdale, NJ: Erlbaum.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension.* New York, NY: Academic Press.

Whitney, P., Ritchie, B., & Clark, M. (1991). Working memory capacity and the use of elaborative inferences in text comprehension. *Discourse Processes*, *14*, 133–145. doi:10.1080/01638539109544779

Wiley, J., & Myers, J. L. (2003). Availability and accessibility of information and causal inferences from scientific text. *Discourse Processes*, *36*, 109–129. doi:10.1207/ S15326950DP3602_2

Williams, J. P., Taylor, M. B., & de Cani, J. S. (1984). Constructing macrostructure for expository text. *Journal of Educational Psychology*, *76*, 1065–1075. doi:10.1037/ 0022-0663.76.6.1065

Wolfe, M. B. (2005). Memory for narrative and expository text: Independent influences of semantic associations and text organization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 359–364. doi:10.1037/0278-7393.31.2.359

Wolfe, M. B., & Mienko, J. A. (2007). Learning and memory of factual content from narrative and expository text. *British Journal of Educational Psychology*, *77*, 541–564. doi:10.1348/000709906X143902

Wolfe, M. B., & Woodwyk, J. M. (2010). Processing and memory of information presented in

narrative or expository texts. *British Journal of Educational Psychology, 80*, 341–362. doi:10.1348/000709910X485700

Yoshida, M. (2003). Working memory capacity and the use of inference in L2 reading. *JACET Bulletin, 36*, 1–17.

Yoshida, M. (2008). Think-aloud protocols and type of reading task: The issue of reactivity in L2 reading research. In M. Bowles, R. Foote, S. Perpiňán, & R. Bhatt (Eds.), *Selected proceedings of the 2007 second language research forum* (pp. 199–209). Somerville, MA: Cascadilla Proceedings Project.

Yoshida, M. (2012). The interplay of processing task, text type, and proficiency in L2 reading. *Reading in a Foreign Language*, *24*, 1–29.

Zhang, H., & Hoosain, R. (2001). The influence of narrative text characteristics on thematic inference during reading. *Journal of Research in Reading*, *24*, 173–186. doi:10.1111/1467-9817.00140

Zhang, H., & Hoosain, R. (2005). Activation of themes during narrative reading. *Discourse Processes*, *40*, 57–82. doi:10.1207/s15326950dp4001_3

Zwaan, R. A. (1994). Effect of genre expectations on text comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 920–933.

Zwaan, R. A. (1996). Processing narrative time shifts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1196–1207. doi:10.1037/0278-7393.22.5.1196

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science, 5*, 292–297. doi:10.1111/j.1467-9280.1995.tb00513.x

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123*, 162–185. doi:10.1037/0033-2909.123.2.162

Zwaan, R. A., Radvansky, A., & Whitten, S. N. (2002). Situation models and themes. In M.

Louwerse & W. van Peer (Eds.), *Thematics: Interdisciplinary studies* (pp. 35–53). Amsterdam, Netherlands: John Benjamins.

Appendices

Appendix 1

Experimental Passages in Experiments 1, 2, and 3

The underlined sentences were only presented in the explicit version.

1. Alice (original title: *Using an elephant gun to kill a fly*)

One morning, Alice discovered *acne on her nose. She had not had acne in years and disliked her skin condition. Alice decided to fly to a famous clinic in Canada for treatment. After the operation, the doctor instructed her not to smile for over two weeks. Alice was happy that the acne had disappeared completely. However, she would have to work very hard to pay off her expensive medical bills. Alice should not have made a simple problem more difficult than it already was.

*acne にきび

2. Bill (original title: *Cutting off your nose to spite your face*)

Bill's teacher was very strict and insisted that the students could use their photography darkroom only a few hours a day. Bill disagreed with the rule because he loved photography. Bill was angry with the school, and he planned to ruin all the chemicals in the darkroom. He mixed all the photographic chemicals together and wasted them. As a result, the darkroom ran out of chemicals. Bill had to wait for more than two months for new chemicals to arrive at the school. Bill made the situation worse for himself with his careless actions when he was angry.

3. Brown (original title: *Too many cooks spoil the broth*)

Mr. and Mrs. Brown wanted to attend a house party. Mr. Brown hired 15-year-old twins as

babysitters for his daughter. He thought the two girls would work together better than if he had hired only one babysitter. However, the twins fought most of the evening about who should do the work. When Mr. Brown came home at midnight, his daughter was still awake. She was crying, while the twins were sleeping soundly. Twins do not always do a job better than one person does.

4. Burt (original title: *Every cloud has a silver lining*)

Burt put in long hours as a night security officer. One day, a large box accidentally fell on him and broke his shoulder. Burt was in a pretty bad condition. He had to spend several months at home recovering his strength. While at home, he started reading about *electronics and decided to take courses by mail. By the time he recovered, Burt had *qualified for and found a better-paying job in an electronics repair shop. He also had more time to enjoy himself. Good things happened to Burt even in a situation that seemed difficult or bad.

*electronics: 電子工学　*qualified for: 資格をとる

5. Ernie (original title: *Counting your chickens before they're hatched*)

Ernie was really encouraged about his interview for a security officer at the new factory in town. The interview was long, and Ernie thought he had done well. He was sure that he would soon be employed as a security officer. He went to the shopping mall because he wanted a dark blue security guard uniform, and finally bought several. The next day he received a phone call from the factory manager saying he was not selected for a security guard position. Ernie was disappointed that he had wasted money on uniforms. Ernie should not make clear plans for something that has not yet happened.

6. Joe (original title: *The blind leading the blind*)

Joe was worried that his business was failing. Joe wanted to ask someone for advice. He talked to his friend Nancy instead of talking to a banker or another store owner. Nancy used to own a restaurant in the next town, but her business had failed because of poor profits. Nancy told Joe he should increase the price of goods so that he would make more profit on each item. Joe took Nancy's advice, but then his sales suddenly fell, and Joe failed in his business. Joe should have taken advice from people with more knowledge or experience about business.

7. Karen (original title: *The pot calling the kettle "black"*)

Karen's swimming coach was a very strict person. He had the team working out for many hours every day. Besides their workouts, he insisted that each player be in the best physical condition for the season. During the pre-season training sessions, the coach would warn the players that they should avoid drinking, drugs, overeating, and especially smoking. "Everyone knows that athletes should treat their bodies with respect," the coach said as he puffed heavily on his long cigarette. The coach should not criticize the players because he was doing the same thing as them.

8. Phil (original title: *Closing the barn door after the horse is gone*)

Phil was in love with his secretary and was well aware that she wanted to marry him. However, Phil was afraid of responsibility, so he kept dating others and delayed proposing to her. Finally, his secretary got tired, began dating, and fell in love with an *accountant. When Phil found out, he went to her and proposed marriage, showing her the ring he had bought. But by that time, his secretary was already planning her honeymoon with the accountant. It was too late for Phil to prevent something that he had noticed.

*accountant: 会計士

Appendix 2

Statements for the Inference Verification Task in Experiment 1

1. Alice

Theme		Alice は単純な問題を難しくするべきではなかった。
Explicit	Goal	Alice は病院に行こうと思った。
	Action	Alice はにきびを見つけた。
	Emotion	Alice はにきびがなくなって喜んだ。
	State	治療費はとても高かった。
Inference	Goal	Alice はにきびをなくしたかった。
	Action	Alice は医者の診察を受けた。
	Emotion	Alice はいやな気分だった。
	State	Alice は外見をとても気にする人だ。
Inappropriate	Goal	Alice は大学に行きたかった。
	Action	Alice はたくさんの本を読んだ。
	Emotion	医者はとても怒っていた。
	State	病院は Alice の家から近かった。
	Theme	Alice には他人を批判する資格がない。

2. Bill

Theme		Bill は自分自身で状況を悪くしてしまった。
Explicit	Goal	Bill は薬品をだめにしてしまおうとした。
	Action	Bill は薬品をまぜた。
	Emotion	Bill は学校に対して怒っていた。
	State	先生はとてもきびしかった。
Inference	Goal	Bill は暗室をもっと使いたかった。
	Action	Bill は学校の暗室に行った。
	Emotion	Bill はとてもくやしかった。
	State	Bill は学生であった。
Inappropriate	Goal	Bill はテスト勉強をしたかった。
	Action	Bill は友人と遊んでいた。
	Emotion	Bill はとても喜んでいた。
	State	Bill は会社員であった。
	Theme	Bill の行いは、良い方向に転じた。

3. Brown

Theme		ふたごの方が 1 人より良いとは限らない。
Explicit	Goal	Brown 夫妻はパーティーに行きたかった。
	Action	ふたごはケンカをした。
	Emotion	Brown さんの娘は泣いていた。
	State	ふたごは 15 才だった。
Inference	Goal	Brown さんは娘を世話してほしかった。
	Action	ふたごは Brown 夫妻の家にやってきた。
	Emotion	Brown さんはあきれた。
	State	ふたごは仲が悪かった。
Inappropriate	Goal	Brown さんは家で休みたかった。
	Action	Brown さんは子どもと遊んでいた。
	Emotion	Brown さんの娘は喜んだ。
	State	ふたごはとても年老いていた。
	Theme	Brown さんはもっと早く決断すべきだった。

4. Burt

Theme		Burt に悪いことが起きたが、良い方向に転じた。
Explicit	Goal	Burt はコースを受けようと思った。
	Action	Burt は修理店で働きはじめた。
	Emotion	Burt はひどく落ち込んだ。
	State	Burt は警備員だった。
Inference	Goal	Burt は資格がほしかった。
	Action	Burt は警備員の仕事をやめた。
	Emotion	Burt はとても嬉しかった。
	State	Burt は前向きな人だった。
Inappropriate	Goal	Burt は外で遊びたかった。
	Action	Burt は旅行に行った。
	Emotion	Burt はとても怒っていた。
	State	Burt は学校の先生だった。
	Theme	Burt が気付いた時には、もう遅かった。

5. Ernie

Theme		Ernie は小さな問題を大げさにするべきではなかった。
Explicit	Goal	Ernie は警備服が欲しかった。
	Action	Ernie はショッピングモールに行った。
	Emotion	Ernie は落ち込んだ。
	State	面接の時間は長かった。
Inference	Goal	Ernie は警備員になりたかった。
	Action	Ernie は求人に申し込んだ。
	Emotion	Ernie はとてもわくわくしていた。
	State	Ernie は自信家であった。
Inappropriate	Goal	Ernie は友達に会いたかった。
	Action	Ernie は散歩に出かけた。
	Emotion	Ernie はとてもおびえていた。
	State	Ernie は面接官であった。
	Theme	Ernie は小さな問題を大げさにするべきではなかった。

6. Joe

Theme		Joe はもっと知識がある人から助言をもらうべきだった。
Explicit	Goal	Joe はアドバイスがほしかった。
	Action	Joe は Nancy に相談をした。
	Emotion	Joe は不安な気持ちだった。
	State	Nancy は以前経営者だった。
Inference	Goal	Joe は経営を良くしたかった。
	Action	Joe は商品の値上げをした。
	Emotion	Joe はとてもがっかりした。
	State	Joe は他人を信用しやすい。
Inappropriate	Goal	Joe は店員を増やしたかった。
	Action	Nancy は Joe に友人を紹介した。
	Emotion	Nancy は急に怒り出した。
	State	Joe は学生であった。
	Theme	Joe はまだ決まっていない計画を立てるべきでなかった。

7. Karen

Theme		コーチは自分ができないことを他人に言うべきではない。
Explicit	Goal	コーチは選手にたくさん練習させたかった。
	Action	コーチは選手に注意をした。
	Emotion	コーチは喫煙者だった。
	State	コーチはとてもきびしかった。
Inference	Goal	コーチは選手を強くしたかった。
	Action	コーチは選手の健康を管理した。
	Emotion	選手たちはコーチにあきれた。
	State	コーチは自分に甘かった。
Inappropriate	Goal	コーチはレストランへ行きたかった。
	Action	コーチは街へ買い物に行った。
	Emotion	コーチは突然悲しくなった。
	State	水泳の練習時間は短かった。
	Theme	コーチは多い方が良いとはかぎらない。

8. Phil

Theme		Phil が気付いた時には手遅れだった。
Explicit	Goal	秘書は Phil と結婚したかった。
	Action	Phil は指輪を買った。
	Emotion	Phil の秘書は疲れてしまった。
	State	Phil と秘書は恋をしていた。
Inference	Goal	Phil は秘書をとりもどそうとした。
	Action	秘書は会計士と結婚した。
	Emotion	Phil はくやしい気持ちだった。
	State	Phil は優柔不断であった。
Inappropriate	Goal	Phil は会計士と仲良くなりたかった。
	Action	会計士は買い物に出かけた。
	Emotion	会計士はとても悲しかった。
	State	秘書はお金持ちだった。
	Theme	Phil が行動を起こすのは早すぎた。

Appendix 3

Examples of Participants' Answers for the Thematic Inference Task in Experiment 2

Original answers were written in Japanese and translated in English by the author.

Burt

Category	Examples
Correct	There is something good even in a situation that seems difficult or bad.
Incorrect	You should be careful not to get hurt while you are working. (Narrow)
	You never know what will happen in your life. (Broad)

Ernie

Category	Examples
Correct	You should not make clear plans for something that has not occurred yet.
Incorrect	Job interviews require careful preparation. (Narrow)
	Life often doesn't go as you wish. (Broad)

Karen

Category	Examples
Correct	You should not criticize someone for a fault that you also have.
Incorrect	Health care is the most important thing for athletes. (Narrow)
	You should not be too strict to others. (Broad)

Appendix 4

Examples of Think-Aloud Protocols in Experiment 3

Process level	Category	Example
Analysis	Word analysis Sentence analysis	responsible…責任？ …responsibility／date, date…dating…keep~ingだから、し続ける。
Paraphrase		(Phil was in love with his secretary and was well aware that she wanted to marry him.) お互いに好きだ、と。
Inference	Backward	(his secretary was already planning her honeymoon with the accountant) あー…断られたのね。
	Predictive	この二人は上手くいかない、いかなくなるような気がします。
	Thematic	手遅れになる前に、行動をするべきだった、ってことかな。
Reader response	Association	honeymoon…リゾート？とかに行った？
	Evaluation	ふーんまあ先に、Philが浮気したんだから悪いかなと思います。
	Reaction	かわいそうというか、まあ、あーあというか。
Self-monitoring		Philが何をしてるかがちょっとわかんないけど／何が言いたい文章なんだろうな…

Note. The verbal data were mostly reported in Japanese.

Appendix 5

Experimental Passages in Experiments 4, 5, and 6

The underlined sentences were only presented in the explicit version.

Africa's Great Green Wall

The Sahara Desert, in Africa, is the world's largest hot desert, and it is getting bigger all the time. Researchers say that, partly because of global warming, the desert is now spreading southward by up to about 50 kilometers a year. This has made life very hard for people in the countries that are directly south of the Sahara. Now, however, a major plan to stop the desert's growth is about to be put into practice.

This plan is known as the Great Green Wall, and it involves the creation of a "wall" of trees 15 kilometers wide and almost 8,000 kilometers long. The Great Green Wall is intended to reduce damage from the sandstorms blowing off the Sahara and to help keep the soil stable and fertile. Its trees will provide local people with wood and other materials, and it will become a home for plants and animals. The trees will also help to remove carbon dioxide from the air. Eleven different countries across Africa have agreed to participate in the project.

The idea itself is not a new one. In fact, it was first suggested in the 1980s. The problem has been a shortage of money. This all changed in 2011, when a group of international organizations, including the Global Environment Facility (GEF), agreed to donate up to $3 billion to the project. A number of NGOs representing local communities, however, are concerned about the plan's possible effects. They are especially worried about the idea of planting a large number of trees. They say this might mean introducing new kinds of trees from other areas that would damage local ecosystems and use up valuable farmland.

The GEF, however, says that these fears are unnecessary. They say that they are not just planning to have trees planted across Africa. Rather, they will require each country to consult with local citizens and come up with a plan that will improve people's lives without harming the environment. As the GEF points out, simply planting trees will not work unless local people have some reason to look after them. By choosing trees that offer a source of income, such as fruit trees, the GEF believes that the Great Green Wall will continue to help people long after it has been completed. <u>The Great Green Wall has many good points for stopping the desert's growth.</u>

Natural Solutions

Malaria is a serious disease that affects millions of people every year. Malaria is spread by female mosquitoes, and one obvious way to fight it is to reduce the number of mosquitoes. This can be done very effectively by using chemicals that kill them. In fact, since the 1950s, this method has resulted in a large reduction in the number of malaria cases. The chemicals used to kill mosquitoes, however, have various disadvantages. Not only are they expensive, but they are often bad for the environment. Moreover, over time, mosquitoes gradually stop being affected by them.

For these reasons, scientists have recently been looking at alternative methods of controlling mosquitoes. One of these is using fish. Mosquitoes lay their eggs in water, and the eggs then turn into tiny worms that live in the water for one or two weeks. Some kinds of fish eat these worms, so introducing these fish into lakes and ponds can lead to fewer mosquitoes. Projects carried out in India have found that, depending on the kind of fish, this method can reduce the number of mosquitoes by over 90 percent.

Using fish to control mosquitoes has many advantages. One of these is that it does no damage to the environment. Another is that the fish reproduce by themselves, making this a very cheap method of fighting malaria. In addition, some of these fish can be caught and sold, meaning that local people can actually earn an income from them. All of these factors are especially important in the developing countries where malaria is still common.

Using nature to control nature in this way is known as "biocontrol." Biocontrol itself is not a new idea, but scientists are now doing more research on it. One problem with using fish has been that it must be limited to permanent bodies of water, such as lakes. Mosquitoes, though, often lay their eggs in pools of rainwater that later dry up. Scientists have now found a kind of fish in the African country of Tanzania that can survive even when these pools are dry. When the rain comes, the fish eat the mosquitoes. Many experts believe that this kind of research is giving biocontrol a bright future as a way to fight diseases. Using fish to control mosquitoes is an effective way of reducing malaria.

Appendix 6

Target Statements (the NS Text) in the Inference Verification Task in Experiment 4

(a) Consistent	魚を使って蚊を減らすことは、マラリア対策として有効である。 [Using fish to control mosquitoes is an effective way of reducing malaria.]
(b) Inconsistent	魚を使って蚊を減らす方法が使える場所は、限られている。 [Using fish to control mosquitoes can be adopted in limited places.]
(c) Off-topic	海に生息している魚の数は、年々減ってきている。 [The number of fish living in the sea is decreasing every year.]
(d) Explicit	マラリアは、メスの蚊によって広められる病気である。 [Malaria is a disease spread by female mosquitoes.] インドで行ったプロジェクトでは、蚊の数が大きく減った。 [Projects carried out in India found that the number of mosquitoes became small.] タンザニアには、水なしでも生きられる魚がいる。 [A kind of fish in Tanzania can survive without water.]
(e) Inappropriate	蚊を殺す化学薬品は、値段がとても安かった。 [The chemicals used to kill mosquitoes are very cheap.] バイオコントロールは、最近提唱されたアイディアである。 [Biocontrol is a new idea that is proposed in recent years.] 蚊の幼虫は、生後 1~2 週間は地上で生活をする。 [The worms of mosquitoes live on the ground for one or two weeks.]

Appendix 7

Examples of Think-Aloud Comments in Experiment 6

Process level	Category	Example
Analysis	Word analysis Sentence analysis	disadvantage… “dis” だから、なんか、「できない」のかな。
Paraphrase		(*this method can reduce the number of mosquitoes by over 90 percent*) 蚊の数は 10%に減った。
Inference	Backward	蚊の幼虫だから、ボウフラか…。
	Predictive	結果的に、これは地球温暖化対策になるんだろうな。
Reader response	Association	サハラの南はたしか土壌が褐色土だったし…。
	Evaluation	この計画は、本当に実現できるのだろうか？／これはとても良い計画だと思います。
	Reaction	とても面白いと思います。
Self-monitoring		この文の意味がちょっとわからないので、次に進みます。
Other		

Note. The verbal data were mostly reported in Japanese.